THE KEY TO EVERYTHING

XXXX
The Original
MOONSHINE
Clear Corn Whiskey
Handcrafted in Small Batches
DISTILLED FOUR TIMES TO REACH PERFECTION
100% CORN WHISKEY
DISTILLED AND BOTTLED BY
STILLHOUSE DISTILLERY
CULPEPER • VIRGINIA • USA
ALCOHOL BY VOL. (80 PROOF) • 750 ML
STILLHOUSE

arXiv:0706.3359v1 [hep-th] 22 Jun 2007

hep-th/yymmnnn

Three-Dimensional Gravity Reconsidered

Edward Witten

School of Natural Sciences, Institute for Advanced Study
Princeton, New Jersey 08540

We consider the problem of identifying the CFT's that may be dual to pure gravity in three dimensions with negative cosmological constant. The c-theorem indicates that three-dimensional pure gravity is consistent only at certain values of the coupling constant, and the relation to Chern-Simons gauge theory hints that these may be the values at which the dual CFT can be holomorphically factorized. If so, and one takes at face value the minimum mass of a BTZ black hole, then the energy spectrum of three-dimensional gravity with negative cosmological constant can be determined exactly. At the most negative possible value of the cosmological constant, the dual CFT is very likely the monster theory of Frenkel, Lepowsky, and Meurman. The monster theory may be the first in a discrete series of CFT's that are dual to three-dimensional gravity. The partition function of the second theory in the sequence can be determined on a hyperelliptic Riemann surface of any genus. We also make a similar analysis of supergravity.

June, 2007

Contents

1. Introduction . . . 1
  1.1. Relation To Gauge Theory . . . 2
  1.2. What To Aim For . . . 5
  1.3. A Non-Classical Restriction . . . 6
  1.4. Plan Of This Paper . . . 7
2. Gauge Theory And The Value Of c . . . 9
  2.1. Quantization Of Parameters . . . 11
  2.2. Comparison To Three-Dimensional Gravity . . . 15
  2.3. Holomorphic Factorization . . . 16
  2.4. Interpretation . . . 19
  2.5. Analog For Supergravity . . . 25
3. Partition Functions . . . 28
  3.1. The Bosonic Case . . . 29
  3.2. The Supersymmetric Case . . . 38
  3.3. Extremal SCFT's With Small k^* . . . 47
4. $k = 2$ Partition Function On A Hyperelliptic Riemann Surface . . . 58
  4.1. Twist Fields . . . 59
  4.2. The $\mathcal{E} \cdot \mathcal{E}$ Operator Product . . . 63
  4.3. Determining The Four-Point Function . . . 66
  4.4. The Genus One Partition Function Revisited . . . 70
  4.5. Role Of The Conformal Anomaly . . . 72
Appendix A. More Extremal Partition Functions . . . 76

1. Introduction

Three-dimensional pure quantum gravity, with the Einstein-Hilbert action

$$I = \frac{1}{16\pi G}\int d^3x\sqrt{g}\left(R + \frac{2}{\ell^2}\right), \tag{1.1}$$

has been studied from many points of view (see [1-3] for some early developments and [4] for a recent review and references), but its status is fundamentally unclear. The present paper is devoted to a tentative attempt to reconsider it. Before giving an overview of the paper, we begin with an introduction to the problem.

The first thought about this theory is that at the classical level it is "trivial," in the sense that there are no gravitational waves, and any two solutions are equivalent locally. So perhaps it might be tractable quantum mechanically.

A second thought is that despite being "trivial," the theory actually is unrenormalizable by power counting, since the gravitational constant G has dimensions of length. So perhaps the quantum theory does not exist.

The claim about unrenormalizability, however, is fallacious, precisely because the classical theory is trivial. In three dimensions, the Riemann tensor R_{ijkl} can be expressed in terms of the Ricci tensor R_{ij}. In the case of pure gravity, the equations of motion express the Ricci tensor as a constant times the metric. So finally, any possible counterterm can be reduced to a multiple of $\int d^3x\sqrt{g}$ and is equivalent on-shell to a renormalization of the cosmological constant, which is parametrized in (1.1) via the parameter ℓ^2. A counterterm that vanishes on shell can be removed by a local redefinition of the metric tensor g (of the general form $g_{ij} \to g_{ij} + aR_{ij} + \ldots$, where a is a constant and the ellipses refer to local terms of higher order). So a more precise statement is that any divergences in perturbation theory can be removed by a field redefinition and a renormalization of ℓ^2.

1.1. Relation To Gauge Theory

The claim just made is valid regardless of how one formulates perturbation theory. But actually, there is a natural formulation in which no field redefinition or renormalization is needed. This comes from the fact that classically, $2+1$-dimensional pure gravity can be expressed in terms of gauge theory. The spin connection ω is an $SO(2,1)$ gauge field (or an $SO(3)$ gauge field in the case of Euclidean signature). It can be combined with the "vierbein" e to make a gauge field of the group $SO(2,2)$ if the cosmological constant is negative (and a similar group if the cosmological constant is zero or positive). We simply combine ω and e to a 4×4 matrix A of one-forms:

$$A = \begin{pmatrix} \omega & e/\ell \\ -e/\ell & 0 \end{pmatrix}. \tag{1.2}$$

Here ω fills out a 3×3 block, while e occupies the last row and column. As long as e is invertible, the usual transformations of e and ω under infinitesimal local Lorentz transformations and diffeomorphisms combine together into gauge transformations of A. This statement actually has a close analog in any spacetime dimension d, with $SO(d-1,2)$ replacing $SO(2,2)$. What is special in $d=3$ is that [5,6] it is also possible to write the action in a gauge-invariant form. Indeed the usual Einstein-Hilbert action (1.1) is equivalent to a Chern-Simons Lagrangian[1] for the gauge field A:

$$I = \frac{k}{4\pi}\int \mathrm{tr}^* \left(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right). \tag{1.3}$$

---

[1] Here tr^* denotes an invariant quadratic form on the Lie algebra of $SO(2,2)$, defined by $\mathrm{tr}^*\, ab = \mathrm{tr}\ a\star b$, where tr is the trace in the four-dimensional representation and $\star$ is the Hodge star, $(\star b)_{ij} = \frac{1}{2}\epsilon_{ijkl}b^{kl}$.

In dimensions other than three, it is not possible to similarly replace the Einstein-Hilbert action with a gauge-invariant action for gauge fields.[2]

In the gauge theory description, perturbation theory is renormalizable by power counting, and is actually finite, because there are no possible local counterterms. The Chern-Simons functional itself is the only gauge-invariant action that can be written in terms of A alone without a metric tensor; as it is not the integral of a gauge-invariant local density, it will not appear as a counterterm in perturbation theory. The cosmological constant cannot be renormalized, since in the gauge theory description, it is a structure constant of the gauge group.

As we have already remarked, the Chern-Simons description of three-dimensional gravity is valid when the vierbein is invertible. This is so for a classical solution, so it is true if one is sufficiently close to a classical solution. Perturbation theory, starting with a classical solution, will not take us out of the region in which the vierbein is invertible, so the Chern-Simons description of three-dimensional gravity is valid perturbatively. The fact that, in this formulation, the perturbation expansion of three-dimensional gravity is actually finite can reasonably be taken as a hint that the corresponding quantum theory really does exist.

However, nonperturbatively, the relation between three-dimensional gravity and Chern-Simons gauge theory is unclear. For one thing, in Chern-Simons theory, nonperturbatively the vierbein may cease to be invertible. For example, there is a classical solution with $A = \omega = e = 0$. The viewpoint in [6] was that such non-geometrical configurations must be included to make sense of three-dimensional quantum gravity nonperturbatively. But it has has been pointed out (notably by N. Seiberg) that when we do know how to make sense of quantum gravity, we take the invertibility of the vierbein seriously. For example, in perturbative string theory, understood as a model of quantum gravity in two spacetime dimensions, the integration over moduli space of Riemann surfaces that leads to a sensible theory is derived assuming that the metric should be non-degenerate.

There are other possible problems in the nonperturbative relation between three-dimensional gravity and Chern-Simons theory. The equivalence between diffeomorphisms and gauge transformations is limited to diffeomorphisms that are continuously connected to the identity. However, in gravity, we believe that more general diffeomorphisms (such as modular transformations in perturbative string theory) play an important role. These are not naturally incorporated in the Chern-Simons description. One can by hand supplement the gauge theory description by imposing invariance under disconnected diffeomorphisms, but it is not clear how natural this is.

---

[2] In $d = 4$, it is possible to write the Hamiltonian constraints of General Relativity in terms of gauge fields [7]. This has been taken as the starting point of loop quantum gravity.

Similarly, in quantum gravity, one expects that it is necessary to sum over the different topologies of spacetime. Nothing in the Chern-Simons description requires us to make such a sum. We can supplement the Chern-Simons action with an instruction to sum over three-manifolds, but it is not clear why we should do this.

From the point of view of the Chern-Simons description, it seems natural to fix a particular Riemann surface Σ, say of genus g, and construct a quantum Hilbert space by quantizing the Chern-Simons gauge fields on Σ. (Indeed, there has been remarkable progress in learning how to do this and to relate the results to Liouville theory [8-11].) In quantum gravity, we expect topology-changing processes, such that it might not be possible to associate a Hilbert space with a particular spatial manifold.

Regardless of one's opinion of questions such as these, there is a more serious problem with the idea that gravity and gauge theory are equivalent non-perturbatively in three dimensions. Some years after the gauge/gravity relation was suggested, it was discovered by Bañados, Teitelboim, and Zanelli [12] that in three-dimensional gravity with negative cosmological constant, there are black hole solutions. The existence of these objects, generally called BTZ black holes, is surprising given that the classical theory is "trivial." Subsequent work [13,14] has made it clear that three-dimensional black holes should be taken seriously, particularly in the context of the AdS/CFT correspondence [15].

The BTZ black hole has a horizon of positive length and a corresponding Bekenstein-Hawking entropy. If, therefore, three-dimensional gravity does correspond to a quantum theory, this theory should have a huge degeneracy of black hole states. It seems unlikely that this degeneracy can be understood in Chern-Simons gauge theory, because this essentially topological theory has too few degrees of freedom. However, some interesting attempts have been made; for a review, see [4].

The existence of the BTZ black hole makes three-dimensional gravity a much more exciting problem. This might be our best chance for a solvable model with quantum black holes. Surely in 3+1 dimensions, the existence of gravitational waves with their nonlinear interactions means that one cannot hope for an exact solution of any system that includes quantum gravity.[3] There might be an exact solution of a 1+1-dimensional model with black holes (interesting attempts have been made [16]), but such a model is likely to be much less realistic than three-dimensional pure gravity. For example, in 1+1 dimensions, the horizon of a black hole just consists of two points, so there is no good analog of the area of the black hole horizon.

1.2. What To Aim For

So we would like to solve three-dimensional pure quantum gravity. But what would it mean to solve it?

First of all, we will only consider the case that the cosmological constant Λ is negative. This is the only case in which we know what it would mean to solve the theory.

Currently, there is some suspicion (for example, see [17]) that quantum gravity with $\Lambda > 0$ does not exist nonperturbatively, in any dimension. One reason is that it does not appear possible with $\Lambda > 0$ to define precise observables, at least none [18] that can be measured by an observer in the spacetime.[4] This is natural if a world with positive Λ (like the one we may be living in) is always at best metastable – as is indeed the case for known embeddings of de Sitter space in string theory [20]. If that is so, then pure gravity with $\Lambda > 0$ does not really make sense as an exact theory in its own right but (like an unstable particle) must be studied as part of a larger system. There may be many choices of the larger system (for example, many embeddings in string theory) and it may be unrealistic to expect any of them to be soluble.

Whether that is the right interpretation or not, we cannot in this paper attempt to solve three-dimensional gravity with $\Lambda > 0$, since, not knowing how to define any mathematically precise observables, we do not know what to try to calculate.

For $\Lambda = 0$, above three dimensions there is a precise observable in quantum gravity, the S-matrix. However, in the three-dimensional case, there is no S-matrix in the usual sense, since in any state with nonzero energy, the spacetime is only locally asymptotic to Minkowski space at infinity [3]. More relevantly for our purposes, in three-dimensional pure gravity with $\Lambda = 0$, there is no S-matrix since there are no particles that can be scattered. There are no gravitons in three dimensions, and there are also no black holes unless $\Lambda < 0$. So again, we do not have a clear picture of what we would aim for to solve three-dimensional gravity with zero cosmological constant.

---

[3] An exact solution or at least an illuminating description of the appropriate Hamiltonian for near-extremal black holes interacting with external massless particles is conceivable.

[4] At least perturbatively, the de Sitter/CFT correspondence gives observables that can be measured by an observer who looks at the whole universe from the outside [19,18]. These observables characterize the wavefunction of the ground state.

With negative cosmological constant, there is an analog, and in fact a much richer analog, of the S-matrix, namely the dual conformal field theory (CFT). This is of course a *two*-dimensional CFT, defined on the asymptotic boundary of spacetime. Not only does AdS/CFT duality make sense in three dimensions, but in fact one of the precursors of the AdS/CFT correspondence was the discovery by Brown and Henneaux [21] of an asymptotic Virasoro algebra in three-dimensional gravity. They considered three-dimensional gravity with negative cosmological constant possibly coupled to additional fields. The action is

$$I = \frac{1}{16\pi G}\int d^3x\sqrt{g}\left(R+\frac{2}{\ell^2}+\ldots\right), \tag{1.4}$$

where the ellipses reflect the contributions of other fields. Their main result is that the physical Hilbert space obtained in quantizing this theory (in an asymptotically Anti de Sitter or AdS spacetime) has an action of left- and right-moving Virasoro algebras with $c_L = c_R = 3\ell/2G$. In our modern understanding [15], this is part of a much richer structure – the boundary conformal field theory.

What it means to solve pure quantum gravity with $\Lambda < 0$ is to find this dual conformal field theory. We focus on the case $\Lambda < 0$ because this is the only case in which we would know what it means to solve the theory. Luckily, and perhaps not coincidentally, this is also the case that has black holes.

1.3. A Non-Classical Restriction

This formulation of what we aim to do makes it clear that we must anticipate a restriction that is rather surprising from a classical point of view. In contemplating the classical action (1.4), it appears that the dimensionless ratio ℓ/G is a free parameter. But the formula for the central charge $c_L = c_R = 3\ell/2G$ shows that this cannot be the case. According to the Zamolodchikov c-theorem [22], in any continuously varying family of conformal field theories in $1+1$ dimensions, the central charge c is constant. More generally, the same is true for the left- and right-moving central charges c_L and c_R.

So the central charges of the dual CFT cannot depend on a continuously variable parameter ℓ/G. It must be [18] that the theory only makes sense for specific values of ℓ/G.

Of course, the c-theorem has an important technical assumption: the theory must have a normalizable and $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$-invariant ground state. (The two factors of $SL(2,\mathbb{R})$ are for left- and right-moving boundary excitations.) This condition is obeyed by three-dimensional gravity, with Anti de Sitter space being the classical approximation to the vacuum. The desired $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ symmetry is simply the classical $SO(2,2)$ symmetry of three-dimensional AdS space.

The statement that ℓ/G cannot be continuously varied is not limited to pure gravity. It holds for the same reason in any theory of three-dimensional gravity plus matter that has a sensible AdS vacuum. For example, in the string theory models whose CFT duals are known, ℓ/G is expressed in terms of integer-valued fluxes; this gives a direct explanation of why it cannot be varied continuously.

1.4. Plan Of This Paper

Now we can describe the plan of this paper.

We aim to solve three-dimensional gravity with negative Λ, at some distinguished values of ℓ/G at which it makes sense.

We do not have any rigorous way to determine the right values. However, in section 2, taking at face value the Chern-Simons description of three-dimensional gravity, we will use it to motivate certain values of ℓ/G. The values that emerge – with the help of a small sleight of hand in the choice of the gauge group for the Chern-Simons theory – are interesting. They are the values at which c_L and c_R are integer multiples of 24, and complete holomorphic factorization of the dual CFT is conceivable.

Not wishing to look a gift horse in the mouth, we will assume that these are the right values to consider. Relying on holomorphic factorization, to describe the solution of the theory, we must describe a sequence of holomorphic CFT's with $c = 24k$, $k = 1, 2, 3, \ldots$.

For $k = 1$, it is believed [23] that there are precisely 71 holomorphic CFT's with the relevant central charge $c = 24$. Of these theories, 70 have some form of Kac-Moody or current algebra symmetry extending the conformal symmetry. In the AdS/CFT correspondence, these theories are dual to three-dimensional theories describing gravity plus additional gauge fields (with Chern-Simons interactions). To describe pure gravity, we need a holomorphic CFT with $c = 24$ and no Kac-Moody symmetry.

Such a model was constructed nearly twenty-five years ago by Frenkel, Lepowsky, and Meurman [24], who also conjectured its uniqueness. The motivation for constructing the model was that it admits as a group of symmetries the Fischer-Griess monster group $\mathbb{M}$ –

the largest of the sporadic finite groups. (The link between the monster and conformal field theory was suggested by developments springing from an observation by McKay relating the monster to the j-function, as we explain more fully in section 3.1.) Arguably, the FLM model is the most natural known structure with $\mathbb{M}$ symmetry. A detailed and elegant description is in the book [25]; for short summaries, see [26,27], and for subsequent developments and surveys, see [28-30]. Assuming the (unproved) uniqueness conjecture of FLM, we propose that their model must give the CFT that is dual to three-dimensional gravity at $c = 24$.

For $c = 24k$, $k > 1$, we need an analog of requiring that there is no Kac-Moody symmetry. A plausible analog, expressing the idea that we aim to describe pure gravity, is that there should be no primary fields of low dimension other than the identity. A small calculation shows that at $c = 24k$, the lowest dimension of a primary other than the identity cannot be greater than $k + 1$, and if we assume that this dimension is precisely $k + 1$, then the partition function is uniquely determined. Conformal field theories with this property were first investigated by Höhn in [31,32] and have been called extremal CFT's; see also [33].

It is not known[5] if extremal CFT's exist for $k > 1$. If such a CFT does exist, it is an attractive candidate for the dual of three-dimensional gravity at the appropriate value of the cosmological constant. The primaries of dimension $k + 1$ and above would be interpreted as operators that create black holes. The dimension $k + 1$ agrees well with the minimum mass of a BTZ black hole. This statement may sound like magic, since the value $k + 1$ is determined from holomorphy and modular invariance without mentioning black holes; but the result is not so surprising if one is familiar with previous results on the AdS/CFT correspondence in three dimensions [34].

Section 3 of this paper is devoted to describing the partition function of an extremal CFT and discussing how such a theory could be related to three-dimensional gravity. In both sections 2 and 3, we consider also the case of three-dimensional supergravity. More precisely, we consider only minimal supergravity, corresponding to $N = 1$ superconformal symmetry for the boundary CFT. In this case, holomorphic factorization is conceivable at $c = 12k^*$, $k^* = 1, 2, 3, \ldots$. Here there is a little ambiguity about exactly what we should mean by an extremal superconformal field theory (SCFT), but pragmatically, there are good candidates at $k^* = 1, 2$. The $k^* = 1$ theory was constructed by Frenkel, Lepowsky, and Meurman, who also conjectured its uniqueness. Its discrete symmetries were understood only recently in work by Duncan [35]. For $k^* = 2$, the extremal SCFT was constructed by Dixon, Ginsparg, and Harvey [26], by modifying the orbifold projection that had been used [24] in constructing the $k = 1$ extremal CFT. Interestingly, the extremal SCFT's with $k^* = 1, 2$ both admit[6] an action of very large discrete groups related to the Conway group. This is a further indication that unusual discrete groups are relevant to three-dimensional gravity and supergravity. In fact, we find some hints that supergravity may have monster symmetry at $k^* = 4$ and baby monster symmetry at $k^* = 6$.

---

[5] Höhn's definition of an extremal CFT allowed holomorphic factorization up to a phase, so that c may be a multiple of 8, not 24. As a result, he discussed several examples of extremal theories that we will not consider here. These examples have k non-integral and less than 2.

Regrettably, we do not know how to construct new examples of extremal conformal or superconformal field theories. Section 4 is devoted to a calculation that aims to give modest support to the idea that new extremal theories do exist. We consider an extremal CFT with $k = 2$ and show that its partition function can be uniquely determined on a hyperelliptic Riemann surface of any genus (including, for example, any Riemann surface of genus 2). The fact that a partition function with the right properties exists and is unique for any genus is hopefully a hint that an extremal $k = 2$ CFT does exist.

We make at each stage the most optimistic possible assumption. Decisive arguments in favor of the proposals made here are still lacking. The literature on three-dimensional gravity is filled with claims (including some by the present author [6]) that in hindsight seem less than fully satisfactory. Hopefully, future work will clarify things.

Advice by J. Maldacena was essential at the outset of this work. I also wish to thank J. Duncan, G. Höhn, G. Nebe, and J. Teschner for descriptions of their work and helpful advice; T. Gannon, R. Griess, J. Lepowsky, and A. Ryba for correspondence about the monster group and related matters; and many colleagues at the IAS and elsewhere, especially A. Maloney, G. Moore, and S. Shenker, for helpful comments.

2. Gauge Theory And The Value Of c

The goal of the present section is to determine what values of the cosmological constant, or equivalently of the central charge c of the boundary CFT, are suggested by the relation between three-dimensional gravity and Chern-Simons gauge theory.

---

[6] The fact that they have essentially the same symmetry group was pointed out by J. Duncan, who also suggested the identification of the $k^* = 2$ theory.

Before proceeding to any calculation, we will dispose of a few preliminary points. The first is that [36], as long as the three-dimensional spacetime is oriented, as we will assume in this paper, three-dimensional gravity can be generalized to include an additional interaction, the Chern-Simons functional of the spin connection ω:

$$\Delta_0 I = \frac{k'}{4\pi}\int_W \mathrm{tr}\left(\omega\wedge d\omega + \frac{2}{3}\omega\wedge\omega\wedge\omega\right). \tag{2.1}$$

Here we think of ω as an $SO(2,1)$ gauge field (or an $SO(3)$ gauge field in the case of Euclidean signature). Also, tr is the trace in the three-dimensional representation of $SO(2,1)$, and k' is quantized for topological reasons (the precise normalization depends on some assumptions and is discussed in sections 2.1 and 2.4). Equivalently, instead of ω, we could use the $SO(2,2)$ gauge field A introduced in eqn. (1.2), and add to the action a term of the form

$$\Delta I = \frac{k'}{4\pi}\int_M \mathrm{tr}\left(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right), \tag{2.2}$$

where now tr is the trace in the four-dimensional representation of $SO(2,2)$. Provided that the conventional Einstein action (1.1) is also present, it does not matter which form of the gravitational Chern-Simons interaction we use, since they lead to equivalent theories. If one adds[7] to ω a multiple of e, the Einstein action (1.1) transforms in a way that cancels the e-dependent part of (2.2), reducing it to (2.1) (while modifying the parameters in the Einstein action).

For our purposes, the $SO(2,2)$-invariant form (2.2) is more useful. This way of writing the Chern-Simons functional places it precisely in parallel with the Einstein-Hilbert action, which as in (1.3) can similarly be expressed as a Chern-Simons interaction, defined with a different quadratic form. We will use the fact that all interactions can be written as Chern-Simons interactions to constrain the proper quantization of all dimensionless parameters, including ℓ/G.

We start with the fact that the group $SO(2,2)$ is locally equivalent to $SO(2,1)\times SO(2,1)$. Moreover, we will in performing the computation assume to start with that $SO(2,1)\times SO(2,1)$ is the right global form of the gauge group. (Then we consider covering groups[8] that are only locally isomorphic to $SO(2,1)\times SO(2,1)$.) Thus, by taking suitable linear combinations of ω and e, we will obtain a pair of $SO(2,1)$ gauge fields A_L and A_R. These have Chern-Simons interactions

---

[7] This operation is invariant under diffeomorphisms and local Lorentz transformations because ω and e transform in the same way under local Lorentz transformations – a statement that holds precisely in three spacetime dimensions.

$$I=\frac{k_L}{4\pi}\int \mathrm{tr}\,\left(A_L\wedge dA_L+\frac{2}{3}A_L\wedge A_L\wedge A_L\right)-\frac{k_R}{4\pi}\int \mathrm{tr}\,\left(A_R\wedge dA_R+\frac{2}{3}A_R\wedge A_R\wedge A_R\right). \tag{2.3}$$

Both k_L and k_R are integers for topological reasons, and this will lead to a quantization of the ratio G/ℓ that appears in the Einstein-Hilbert action, as well as the gravitational Chern-Simons coupling (2.2). The minus sign multiplying the last term in (2.3) is convenient; it will ensure that k_L and k_R are both positive.

2.1. Quantization Of Parameters

For completeness, we begin by reviewing the quantization of the Chern-Simons coupling in gauge theory. The basic case to consider is that the gauge group is $U(1)$. The gauge field A is a connection on a complex line bundle $\mathcal{L}$ over a three-manifold W, which for simplicity we will assume to have no boundary. Naively speaking, the Chern-Simons action is

$$I=\frac{k}{2\pi}\int_W A\wedge dA \tag{2.4}$$

with some coefficient k. If the line bundle $\mathcal{L}$ is trivial, then we can interpret A as a one-form, and I is well-defined as a real-valued functional. If this were the general situation, there would be no need to quantize k.

However, in general $\mathcal{L}$ is non-trivial, A has Dirac string singularities, and the formula (2.4) is not really well-defined as written. To do better, we pick a four-manifold M of boundary W and such that $\mathcal{L}$ extends over M. Such an M always exists. Then we pick an extension of $\mathcal{L}$ and A over M, and replace the definition (2.4) with

$$I_M=\frac{k}{2\pi}\int_M F\wedge F, \tag{2.5}$$

where $F=dA$ is the curvature. Now there is no Dirac string singularity, and the definition of I_M makes sense. But I_M does depend on M (and on the chosen extension of $\mathcal{L}$, though we do not indicate this in the notation). To quantify the dependence on M, we consider two different four-manifolds M and M' with boundary W and chosen extensions of $\mathcal{L}$. We can build a four-manifold X with no boundary by gluing together M and M' along W, with opposite orientation for M' so that they fit smoothly along their common boundary. Then we get

---

[8] For an early treatment of coverings in the context of Chern-Simons theory with compact gauge group, see [37]

$$I_M - I_{M'} = \frac{k}{2\pi}\int_X F\wedge F. \tag{2.6}$$

Now, on the closed four-manifold X, the quantity $\int_X F\wedge F/(2\pi)^2$ represents $\int_X c_1(\mathcal{L})^2$ (here c_1 is the first Chern class) and so is an integer. In quantum mechanics, the action function I should be defined modulo 2π (so that $\exp(iI)$, which appears in the path integral, is single-valued). Requiring $I_M - I_{M'}$ to be an integer multiple of 2π, we learn that k must be an integer. This is the quantization of the Chern-Simons coupling for $U(1)$ gauge theory.[9]

Now let us move on to the case of gauge group $SO(2,1)$. The group $SO(2,1)$ is contractible onto its maximal compact subgroup $SO(2)$, which is isomorphic to $U(1)$. So quantization of the Chern-Simons coupling for an $SO(2,1)$ gauge field can be deduced immediately from the result for $U(1)$. Let A be an $SO(2,1)$ gauge field and define the Chern-Simons coupling

$$I = \frac{k}{4\pi}\int_W \mathrm{tr}\left(A\wedge dA + \frac{2}{3}A\wedge A\wedge A\right), \tag{2.7}$$

where tr is the trace in the three-dimensional representation of $SO(2,1)$. Then, in order for I to be part of the action of a quantum theory, k must be an integer. The reason that the factor $1/2\pi$ in (2.5) has been replaced by $1/4\pi$ in (2.7) is simply that, when we identify $U(1)$ with $SO(2)$ and then embed it in $SO(2,1)$, the trace gives a factor of 2.

Coverings

So we have obtained the appropriate quantization of the Chern-Simons coupling for gauge group $SO(2,1)$. However, this is not quite the whole story, because $SO(2,1)$ is not simply-connected. As it is contractible to $SO(2)\cong U(1)$, it has the same fundamental group as $U(1)$, namely $\mathbb{Z}$. Hence it is possible, for every positive integer n, to take an

---

[9] There is a refinement if the three-manifold W is endowed with a spin structure. In this case, k can be a half-integer, as explained in [38]. This refinement is physically realized in the quantum Hall effect with filling fraction 1. That effect can be described by an electromagnetic Chern-Simons coupling with $k = 1/2$; the half-integral value is consistent because the microscopic theory has fermions and so requires a spin structure.

n-fold cover of $SO(2,1)$. The most familiar of these is the two-fold cover, $SL(2,\mathbb{R})$. In addition, $SO(2,1)$ has a simply-connected universal cover.

We want to work out the quantization of the Chern-Simons interaction if $SO(2,1)$ is replaced by one of these covering groups. Again, it is convenient to start with $U(1)$. To say that the gauge group of an abelian gauge theory is $U(1)$ rather than $\mathbb{R}$ means precisely that the possible electric charges form a lattice, generated by a fundamental charge that we call "charge 1." Dually, the magnetic fluxes are quantized, with $\int_C F/2\pi \in \mathbb{Z}$ for any two-cycle C. Replacing $U(1)$ by an n-fold cover means that the electric charges take values in $n^{-1}\mathbb{Z}$, and dually, the magnetic fluxes are divisible by n, $\int_C F/2\pi \in n\mathbb{Z}$. As a result, for a four-manifold X, we have $\int_C F\wedge F/(2\pi)^2 \in n^2\mathbb{Z}$. So, in requiring that the Chern-Simons function (2.4) should be well-defined modulo 2π, we require that $k \in n^{-2}\mathbb{Z}$. This is the appropriate result for an n-fold cover. In the case of the universal cover, with $U(1)$ replaced by $\mathbb{R}$, the magnetic fluxes vanish and there is no topological restriction on k.

These statements carry over immediately to covers of $SO(2,1)$, whose covers are all contractible to corresponding covers of $U(1)$. So for an n-fold cover of $SO(2,1)$, we require

$$k \in n^{-2}\mathbb{Z}, \tag{2.8}$$

and for the universal cover of $SO(2,1)$, k is arbitrary and can vary continuously.

Diagonal Covers

There is more to say, because three-dimensional gravity is actually related to $SO(2,1)\times SO(2,1)$ gauge theory, not just to gauge theory with a single $SO(2,1)$. So we should consider covers of $SO(2,1)\times SO(2,1)$ that do not necessarily come from separate covers of the two factors.

As $SO(2,1)\times SO(2,1)$ is contractible to $SO(2)\times SO(2) = U(1)\times U(1)$, we can proceed by first analyzing the $U(1)\times U(1)$ case. We consider a $U(1)\times U(1)$ gauge theory with gauge fields A, B and a Chern-Simons action

$$I = \frac{k_L}{2\pi}\int_W A\wedge dA - \frac{k_R}{2\pi}\int_W B\wedge dB. \tag{2.9}$$

To define I in the topologically non-trivial case, we pick a four-manifold M over which everything extends and define

$$I_M = \int_M \left(\frac{k_L}{2\pi} F_A\wedge F_A - \frac{k_R}{2\pi} F_B\wedge F_B\right), \tag{2.10}$$

where F_A and F_B are the two curvatures. This is well-defined mod 2π if

$$I_X = \int_X \left(\frac{k_L}{2\pi} F_A \wedge F_A - \frac{k_R}{2\pi} F_B \wedge F_B \right) \tag{2.11}$$

is a multiple of 2π for any $U(1) \times U(1)$ gauge field over a closed four-manifold X.

In $U(1) \times U(1)$ gauge theory, the charge lattice is generated by charges $(1,0)$ and $(0,1)$, and the cohomology classes $x = F_A/2\pi$ and $y = F_B/2\pi$ are integral. For a cover of $U(1) \times U(1)$, we want to extend the charge lattice. To keep things simple, we will consider only the case that will actually be important in our application: a diagonal cover, in which one adds the charge vector $(1/n, 1/n)$ for some integer n. In this case, x and y are still integral, and their difference is divisible by n: $x = y + nz$ where n is an integral class. We have

$$I_X = 2\pi(k_L - k_R) \int_X y^2 + 2\pi k_L \int_X (n^2 z^2 + 2nyz). \tag{2.12}$$

The condition that this is a multiple of 2π for any X and any integral classes y, z is that

$$\begin{aligned} k_L &\in \begin{cases} n^{-1}\mathbb{Z} & \text{if } n \text{ is odd} \\ (2n)^{-1}\mathbb{Z} & \text{if } n \text{ is even} \end{cases} \\ k_L - k_R &\in \mathbb{Z}. \end{aligned} \tag{2.13}$$

These are also the restrictions on k_L and k_R if the gauge group is a diagonal cover of $SO(2,1) \times SO(2,1)$ with action (2.3). For example, the group $SO(2,2)$ is a double cover of $SO(2,1) \times SO(2,1)$, and this cover corresponds to the case $n = 2$ of the above discussion. So for $SO(2,2)$ gauge theory, the appropriate restriction on the Chern-Simons levels is

$$\begin{aligned} k_L &\in \frac{1}{4}\mathbb{Z} \\ k_L - k_R &\in \mathbb{Z}. \end{aligned} \tag{2.14}$$

For a general n-fold diagonal cover, one should use (2.13).

It is also possible to form a "universal diagonal cover," corresponding roughly to the limit $n \to \infty$ in the above formulas. With this gauge group, there is no restriction on k_L, but $k_L - k_R$ is an integer. In terms of three-dimensional gravity, to which we return next, this corresponds to letting ℓ/G be a freely variable parameter, while the gravitational Chern-Simons coupling k' defined in eqn. (2.2) is an integer. As explained in section 1.3, although this is the state of affairs classically in three-dimensional gravity, it cannot be the correct answer quantum mechanically.

2.2. Comparison To Three-Dimensional Gravity

So far, we have understood the appropriate gauge theory normalizations for the Chern-Simons action

$$\begin{aligned} I =& k_L I_L + k_R I_R \\ =& \frac{k_L}{4\pi}\int \mathrm{tr}\left(A_L\wedge dA_L + \frac{2}{3}A_L\wedge A_L\wedge A_L\right) - \frac{k_R}{4\pi}\int \mathrm{tr}\left(A_R\wedge dA_R + \frac{2}{3}A_R\wedge A_R\wedge A_R\right). \end{aligned} \tag{2.15}$$

Our next step will be to express A_L and A_R, which are gauge fields of $SO(2,1)\times SO(2,1)$ (or a covering group) in terms of gravitational variables, and thereby determine the constraints on the gravitational couplings. We have

$$I = \frac{k_L + k_R}{2}\left(I_L - I_R\right) + (k_L - k_R)\frac{(I_L + I_R)}{2}. \tag{2.16}$$

The term in (2.16) proportional to $I_L - I_R$ will gave the Einstein-Hilbert action (1.1), while the term proportional to $(I_L + I_R)/2$ is equivalent to the gravitational Chern-Simons coupling (2.2) with coefficient $k' = k_L - k_R$.

The spin connection $\omega^{ab} = \sum_i dx^i \omega_i^{ab}$ is a one-form with values in antisymmetric 3×3 matrices. The vierbein is conventionally a one-form valued in Lorentz vectors, $e^a = \sum_i dx^i\, e_i^a$. The metric is expressed in terms of e in the usual way, $g_{ij}dx^i\otimes dx^j = \sum_{ab}\eta_{ab}e^a\otimes e^b$, where $\eta = \mathrm{diag}(-1,1,1)$ is the Lorentz metric; and the Riemannian volume form is $d^3x\sqrt{g} = \frac{1}{6}\epsilon_{abc}e^a\wedge e^b\wedge e^c$, where ϵ_{abc} is the antisymmetric tensor with, say, $\epsilon_{012} = 1$. All this has an obvious analog in any dimension. However, in three dimensions, a Lorentz vector is equivalent to an antisymmetric tensor; this is the fact that makes it possible to relate gravity and gauge theory. It is convenient to introduce ${}^*e_{ab} = \epsilon_{abc}e^c$, which is a one-form valued in antisymmetric matrices, just like ω. We raise and lower local Lorentz indices with the Lorentz metric η, so $\frac{1}{2}\epsilon^{abc}\epsilon_{bcd} = -\delta_d^a$, and $e^c = -\frac{1}{2}\epsilon^{abc}\,{}^*e_{bc}$.

We can combine ω and ${}^*e$ and set $A_L = \omega - {}^*e/\ell$, $A_R = \omega + {}^*e/\ell$. A small computation gives

$$I_L - I_R = -\frac{1}{\pi\ell}\int \mathrm{tr}\,{}^*e(d\omega + \omega\wedge\omega) - \frac{1}{3\pi\ell^3}\int \mathrm{tr}\,({}^*e\wedge{}^*e\wedge{}^*e). \tag{2.17}$$

In terms of the matrix-valued curvature two-form $R^{ab} = (d\omega + \omega\wedge\omega)^{ab} = \frac{1}{2}\sum_{ij} dx^i \wedge dx_j R_{ij}^{ab}$, where R_{ij}^{ab} is the Riemann tensor, and the metric tensor g, this is equivalent to

$$I_L - I_R = \frac{1}{\pi\ell}\int d^3x\sqrt{g}\left(R + \frac{2}{\ell^2}\right). \tag{2.18}$$

Remembering the factor of $(k_L+k_R)/2$ in (2.16), we see that this agrees with the Einstein-Hilbert action (1.1) precisely if

$$k_L+k_R=\frac{\ell}{8G}. \tag{2.19}$$

The central charge of the boundary conformal field theory was originally computed by Brown and Henneaux [21] for the case that the gravitational Chern-Simons coupling $k'=k_L-k_R$ vanishes. In this case, we set $k=k_L=k_R=\ell/16G$. The formula for the central charge is $c=3\ell/2G$, and this leads to $c=24k$. For the case $k'=0$, the boundary CFT is left-right symmetric, with $c_L=c_R$, so in fact $c_L=c_R=24k$.

In general, the boundary CFT has left- and right-moving Virasoro algebras that can be interpreted (for a suitable orientation of the boundary) as boundary excitations associated with A_L and A_R respectively. So the central charges c_L and c_R are functions only of k_L and k_R, respectively. Hence the generalization of the result obtained in the last paragraph is

$$(c_L,c_R)=(24k_L,24k_R). \tag{2.20}$$

2.3. Holomorphic Factorization

In conformal field theory in two dimensions, the ground state energy is $-c/24$. More generally, if there are separate left and right central charges c_L and c_R, the ground state energies for left- and right-movers are $(-c_L/24,-c_R/24)$. Modular invariance says that the difference between the left- and right-moving ground state energies must be an integer. In the above calculation, $(c_L-c_R)/24=k_L-k_R$. According to (2.20), this is an integer provided that the gravitational Chern-Simons coupling $k'=k_L-k_R$ is integral.

Holomorphic factorization requires that the left- and right-moving ground state energies should be separately integral, so that there is modular invariance separately for left-moving and right-moving modes of the CFT. Thus, for holomorphic factorization, c_L and c_R must both be integer multiples of 24. This is the only constraint, since holomorphic CFT's with $c=24$ do exist (and have been classified [23] modulo a conjecture mentioned in section 1.4).

According to (2.20), the condition for c_L and c_R to be multiples of 24 is precisely that k_L and k_R must be integers. As in our discussion of eqn. (2.7), this is the right condition if the gauge group is precisely $SO(2,1)\times SO(2,1)$ rather than a covering group.

It is possible to give an intuitive explanation of why this is the right gauge group if the boundary CFT is supposed to be holomorphically factorized. The Virasoro algebra is, of course, infinite-dimensional, but it has a finite-dimensional subalgebra, generated by $L_{\pm 1}$ and L_0, that is a symmetry of the vacuum. It is customary to refer to the corresponding symmetry group of the vacuum as $SL(2,\mathbb{R})$, but in fact, in a holomorphic CFT, in which all energies are integers, the group that acts faithfully is really $SO(2,1)$. So a holomorphically factorized CFT has symmetry group $SO(2,1)\times SO(2,1)$, and it is natural that this is the right gauge group in a gauge theory description of (aspects of) the dual gravitational theory.

Now let us consider some other possible gauge groups. One possibility is to take a double cover of each factor of $SO(2,1)\times SO(2,1)$, taking the gauge group to be $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$. The appropriate restriction on the gauge theory couplings was determined in (2.8) (where we should set $n=2$) and is that k_L and k_R take values in $\frac{1}{4}\mathbb{Z}$. Hence the central charges c_L and c_R are multiples of 6. In particular, $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$ gauge theory would allow us to consider values of the couplings that contradict modular invariance of the boundary CFT.

There is perhaps a more intuitive argument suggesting that $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$ is not the right group to consider. If the gauge group is $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$, then upon taking the two-dimensional representation of one of the $SL(2,\mathbb{R})$'s, we get a two-dimensional real vector bundle over W which generalizes what in classical geometry is the spin bundle. Thus, this would be a theory in which, in the classical limit, W is a spin manifold, endowed with a distinguished spin structure (or even two of them). That is appropriate in a theory with fermions, but not, presumably, in a theory of pure gravity.

We similarly lose modular invariance and have difficult to interpret geometric structures if we consider other non-diagonal covers of $SO(2,1)\times SO(2,1)$. So let us discuss the diagonal covers of $SO(2,1)\times SO(2,1)$ that were considered at the end of section 2.1. We know that the universal diagonal cover is not right, since then ℓ/G could vary continuously. This leaves the possibility of an n-fold diagonal cover for some n. The best-motivated example is perhaps the two-fold cover $SO(2,2)$, which is the symmetry of Anti de Sitter spacetime (as opposed to a cover of that spacetime). In this case, according to (2.14), k_L and k_R can take values in $\mathbb{Z}/4$, as long as $k'=k_L-k_R$ is integral. For the boundary CFT, this means that c_L and c_R can be multiples of 6 (with their difference a multiple of 24). For example, let us consider a hypothetical CFT of $(c_L,c_R)=(6,6)$. The ground state energies are $(-1/4,-1/4)$. As these values are not integers, the symmetry group of the ground state is not $SO(2,1)\times SO(2,1)$, but a four-fold diagonal cover, which is a double cover of the gauge group $SO(2,2)$ that was assumed. Such a CFT cannot be holomorphically factorized, since the left- and right-moving ground state energies are not integers; it cannot even be holomorphically factorized up to a phase.[10] The structure is considerably more complicated than in the holomorphically factorized case. Similar remarks apply to other diagonal covers.

Pragmatically, the most important argument against trying to describe three-dimensional gravity via a cover of $SO(2,1) \times SO(2,1)$ may be simply the fact that no good candidates are known. For example, no especially interesting bosonic CFT (as opposed to a superconformal field theory) seems to be known at $c = 6, 12$, or 18, which are values that we would expect if we use the gauge group $SO(2,2)$. By contrast, at $c = 24$, which is natural for $SO(2,1) \times SO(2,1)$, the Frenkel-Lepowsky-Meurman monster theory is a distinguished candidate, as we mentioned in section 1.4 and will explain in more detail in section 3.

The simplest possible hypothesis is that the right gauge group to consider in studying three-dimensional pure gravity is $SO(2,1) \times SO(2,1)$, with integral k_L and k_R and holomorphic factorization of the boundary CFT. It would be highly unnatural to overlook the fact that $SO(2,1) \times SO(2,1)$ gauge theory leads to precisely the values of the central charge at which the drastic simplification of the boundary CFT known as holomorphic factorization is conceivable. Moreover, the fact that the classical action (2.15) of the gauge theory is a sum of decoupled actions for A_L and A_R – related respectively to left- and right-moving modes of the boundary CFT – is a hint of holomorphic factorization of the boundary theory.

At any rate, regardless of whether they give the whole story, it does seem well-motivated to look for holomorphically factorized CFT's with central charges multiples of 24 that are dual to three-dimensional pure gravity at special values of the cosmological constant. That will be our main focus in the rest of this paper.

---

[10] The left-moving partition function would have to be $\Phi/\Delta^{1/4}$, where $\Delta = \eta(q)^{24}$ is the discriminant, η being the Dedekind eta function, and Φ is a modular form of weight 3. The power of Δ was determined to get the right ground state energy; modular invariance implies that the modular weight of Φ must equal that of $\Delta^{1/4}$. As there is no modular form of weight 3, such a theory does not exist. Even at $(c_L, c_R) = (12, 12)$, it is not possible to have holomorphic factorization up to a phase. To get modular invariance and the right ground state energy, the left-moving partition function would have to be $E_6/\Delta^{1/2}$, where E_6 is the Eisenstein series of weight 6. However, the coefficients in the q-expansion of this function are not positive.

2.4. Interpretation

A few further words of interpretation seem called for.

We do *not* claim that three-dimensional gravity is equivalent, nonperturbatively, to Chern-Simons gauge theory. Some objections to this idea were described in section 1.1. We know that Chern-Simons gauge theory is useful for perturbation theory, as was explained in that section, and we hope that it is useful for understanding some nonperturbative questions. We used the gauge theory approach to get some hints about the right values of the cosmological constant (or equivalently of the central charge) simply because it was the only tool available. We certainly do not claim to have a solid argument that the values of (c_L, c_R) suggested by $SO(2,1) \times SO(2,1)$ gauge theory are the only relevant ones.

Implicit in the gauge theory approach to quantization of the dimensionless parameter ℓ/G is an assumption that at least some of the non-geometrical states that can be described in gauge theory make sense. Indeed, when understood in classical geometry with the metric assumed to be smooth and nondegenerate, the Einstein-Hilbert action (1.1) is well-defined as a real-valued function. The topological problems that cause it to be multivalued when interpreted in gauge theory as a Chern-Simons function, and lead to quantization of ℓ/G, depend on allowing certain configurations that are natural in gauge theory, but singular in geometry because the vierbein is not invertible.

Furthermore, holomorphic factorization most likely is possible only if the path integral includes a sum over some non-geometrical configurations. In quantum gravity, we at least expect to sum over all topologies of a three-manifold W, perhaps with some fixed asymptotic behavior. The choice of topology should be expected to affect both the left- and right-movers of the boundary CFT. To achieve holomorphic factorization, there must presumably be some sort of separate topological sum for left-movers and right-movers. As this does not occur in classical geometry, it must depend on some sort of non-geometric contributions to the path integral – though these contributions may be exponentially small in many circumstances.

The Gravitational Chern-Simons Action Reconsidered

Our discussion of the quantization of the gravitational Chern-Simons coupling (2.1)

$$\Delta_0 I = \frac{k'}{4\pi} \int_W \operatorname{tr}\left(\omega \wedge d\omega + \frac{2}{3}\omega \wedge \omega \wedge \omega\right) \tag{2.21}$$

has been based entirely on gauge theory. We now want to describe the similar but slightly different answer that would come from classical differential geometry. As always, the differences reflect the fact that the gauge theory analysis allows some non-geometrical configurations.

We first briefly restate the gauge theory analysis. ω is a connection on an $SO(2,1)$ or (in Euclidean signature) $SO(3)$ bundle over a three-manifold W. As usual, to define $\Delta_0 I$ more precisely, we pick an oriented four-manifold M of boundary W with an extension of ω over W. Then we define

$$I_M = \frac{k'}{4\pi}\int_M \mathrm{tr}\, F\wedge F. \tag{2.22}$$

If M is replaced by some other four-manifold M', and $X = M - M'$ is a four-manifold without boundary obtained by gluing together M and M', then

$$I_M - I_{M'} = \frac{k'}{4\pi}\int_X \mathrm{tr}\, F\wedge F = 2\pi k'\int_X p_1(F), \tag{2.23}$$

where $p_1(F) = (1/8\pi^2)\mathrm{tr}\, F\wedge F$ is the first Pontryagin form. In general, $\int_X p_1(F)$ can be any integer, so the condition that the indeterminacy in I_M is an integer multiple of 2π means simply that k' is an integer.

This is the right answer if the right thing to do is to simply think of ω as an $SO(3)$ or $SO(2,1)$ gauge field, ignoring its classical relation to gravity. However, in classical gravity, one can get a better answer. In classical gravity, ω is a connection on the tangent bundle TW of W. Now replace TW by $TW\oplus\epsilon$, where ϵ is a trivial real line bundle. Then ω can be regarded as a connection on $TW\oplus\epsilon$ in an obvious way, and $TW\oplus\epsilon$ extends over M as the tangent bundle of M. With this choice, (2.22) becomes

$$I_M = \frac{k'}{4\pi}\int_M \mathrm{tr}\, R\wedge R, \tag{2.24}$$

where R is the curvature form of M, and (2.23) becomes

$$I_M - I_{M'} = 2\pi k'\int_X p_1(R). \tag{2.25}$$

The effect of this is that instead of the first Pontryagin number of a general bundle over X, as in (2.23), we have here the first Pontryagin number $p_1(TX)$ of the tangent bundle of X. This number is divisible by 3, because of the signature theorem, which says that for a four-manifold X, $p_1(TX)/3$ is an integer, the signature of X. Hence, in the gravitational interpretation, the condition on k' is

$$k' \in \frac{1}{3}\mathbb{Z}. \tag{2.26}$$

There is also a variant of this. If W is a spin manifold and we are willing to define the gravitational Chern-Simons coupling (2.21) in a way that depends on the spin structure of W (this does not seem natural in ordinary gravity, but it may be natural in supergravity, which we come to in section 3.2), the condition on k' can be further relaxed. In this case, we can select M so that the chosen spin structure on W extends over M. If M' is another choice with the same property, then $X = M - M'$ is a spin manifold. But (as follows from the Atiyah-Singer index theorem for the Dirac equation) the signature of a four-dimensional spin manifold is divisible by 16. So in this situation $p_1(TX)$ is a multiple of 48, and the result for k' under these assumptions is

$$k' \in \frac{1}{48}\mathbb{Z}. \tag{2.27}$$

The Physical Hilbert Space

Now we shall discuss the meaning for gravity of the physical Hilbert space of Chern-Simons gauge theory.

In three-dimensional Chern-Simons gauge theory, one can fix a Riemann surface C and construct a Hilbert space $\mathcal{H}_C$ of physical states obtained by quantizing the given theory on C. This Hilbert space depends on the Chern-Simons couplings, so when we want to be more precise, we will call it $\mathcal{H}_C(k_L, k_R)$. In [6], it was proposed that the physical Hilbert space of three-dimensional gravity on a Riemann surface C should be obtained in essentially this way, using $SO(2,1) \times SO(2,1)$ gauge theory (or something similar, depending on whether the cosmological constant is positive, negative, or zero). Some serious objections to the claim that pure gravity and Chern-Simons gauge theory are equivalent in three dimensions were noted in section 1.1. Yet remarkable progress has made made [8,11] in understanding the quantization of $SO(2,1)$ gauge theory, and even more remarkably, in relating this quantization to Liouville theory. It therefore seems likely that the Hilbert space obtained by quantizing Chern-Simons gauge theory of $SO(2,1) \times SO(2,1)$ means something for three-dimensional gravity even if the proposal in [6] was premature. The rest of this subsection will be devoted to an attempt (not used in the rest of the paper) to reconcile the different points of view. See [39] for some useful background.

We may start by asking what we mean by the physical Hilbert space $\mathcal{H}_C$ obtained in quantum gravity by quantizing on a closed manifold C. What type of question is this space supposed to answer? Quantization on, for example, an asymptotically flat spacetime leads to a Hilbert space that can be interpreted in a relatively straightforward way, but the physical meaning of a Hilbert space obtained by quantizing on a compact spatial manifold (if such a Hilbert space can be defined at all) is not clear.

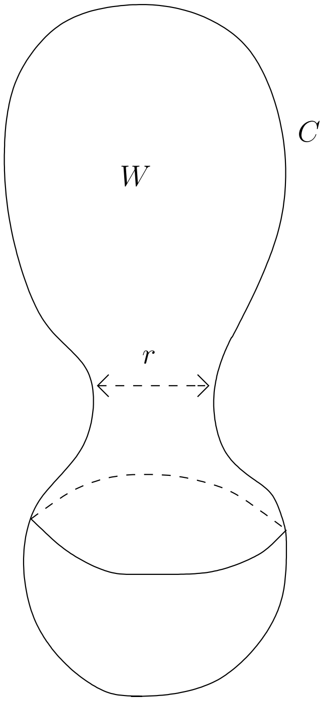


Fig. 1:

The Hartle-Hawking wavefunction Ψ is computed by integrating over three-manifolds W with a given boundary C. The dotted line labeled r denotes a path in W connecting two points in C. It is possible to vary W so that the length of this path goes to zero with no change in the geometry of C. In this limit, W becomes singular even though C is smooth. This gives a "source term" in the Wheeler-de Witt equation, because of which the Hartle-Hawking wavefunction does not obey this equation.

One line of thought that is relatively close to working is to consider the Hartle-Hawking wavefunction [40] and claim that it is a vector in $\mathcal{H}_C$. (The obvious idea that quantum mechanical probabilities would be calculated in terms of inner products of vectors in a Hilbert space $\mathcal{H}_C$ of physical states is afflicted with similar but more serious problems.) The Hartle-Hawking wavefunction is a functional of metrics on C. For every metric h on C, we define $\Psi(h)$ as the result of performing a path integral over three-manifolds W whose boundary is C and whose metric g coincides with h on the boundary. Formally, one can try to argue that $\Psi(h)$ obeys the Wheeler-de Witt equation and thus is a vector in a Hilbert space $\mathcal{H}_C$ of solutions of this equation. Moreover, one can formally match the Wheeler-de Witt equations of gravity with the conditions for a physical state in Chern-Simons gauge theory. Though many steps in these arguments work nicely, one runs into trouble because a Riemann surface can be immersed, rather than embedded, in a three-manifold, and hence it is possible for W to degenerate without C degenerating (fig. 1). As a result, the Hartle-Hawking wavefunction does not obey the Wheeler-de Witt equation and is not a vector in $\mathcal{H}_C$.

In the case of negative cosmological constant, the boundary CFT gives a sort of cure for the problem with the Hartle-Hawking wavefunction. Instead of thinking of C as an ordinary boundary of W, we think of it as a conformal boundary at infinity. The partition function $\widehat{\Psi}(h)$ of the boundary CFT is defined by performing the path integral over all choices of W with C as conformal boundary. This is well-behaved, because, with C at conformal infinity, it is definitely embedded rather than immersed. Moreover, $\widehat{\Psi}(h)$ is a sort of limiting value of the Hartle-Hawking wavefunction. Indeed, let ϕ be a positive function on C. Then $\widehat{\Psi}(h)$ is essentially[11] the limiting value of $\Psi(e^\phi h)$ as $\phi \to \infty$.

This suggests that we should be able to think of $\widehat{\Psi}(h)$ as a vector in the Hilbert space $\mathcal{H}_C$ associated with three-dimensional gravity and a two-manifold C. This viewpoint can be explained more fully. The phase space of $SO(2,1) \times SO(2,1)$ Chern-Simons theory on C is the space of $SO(2,1) \times SO(2,1)$ flat connections on C. The space of $SO(2,1)$ flat connections on C has several topological components (labeled by the first Chern class, which comes in because $SO(2,1)$ is contractible to $SO(2) \cong U(1)$). One of these components, the only one that can be simply interpreted in terms of classical gravity with negative cosmological constant,[12] is isomorphic to Teichmuller space $\mathcal{T}$. Thus, this component of the classical phase space $\mathcal{M}$ is a product of two copies of $\mathcal{T}$, parametrized by a pair of points $\tau, \tau' \in \mathcal{T}$. One can quantize $\mathcal{M}$ (or at least this component of it) naively by using the standard holomorphic structure of $\mathcal{T}$. If we do this, the wavefunction of a physical state is a "function" of τ and τ' that is holomorphic in τ and antiholomorphic in τ'. (Antiholomorphy in one variable reflects the relative minus sign in the Chern-Simons action (2.15); we assume that k_L and k_R are positive.) Actually, a physical state wavefunction $\Psi(\tau, \overline{\tau}')$ is not quite a function of τ and $\overline{\tau}'$ in the usual sense, but is a form, of weights determined by k_L and k_R. So it takes values in a Hilbert space $\mathcal{H}_C(k_L, k_R)$ that depends on the Chern-Simons couplings. Such a wavefunction $\Psi(\tau, \overline{\tau}')$ is determined by its restriction to the diagonal subspace $\tau = \tau'$. Moreover, if we want to make a relation to gravity, it is natural to require that Ψ should be invariant under the diagonal action of the mapping class group on τ and τ'; this condition is compatible with restricting to $\tau = \tau'$.

---

[11] One needs some renormalization in this limit; the necessary renormalization reflects the conformal anomaly. Because of this anomaly, $\widehat{\Psi}(h)$ is not quite a function only of the conformal structure of C but a function of the metric that transforms with a certain weight under conformal rescalings. As a result, the precise choice of ϕ matters, but only in a rather simple way.

[12] In the approach to quantization developed in [8,11], this is the only component considered.

Similarly, the partition function of a CFT on the Riemann surface C is a not necessarily holomorphic "function" (actually a form of appropriate weights) $\Psi(\tau, \overline{\tau})$. Being real analytic, Ψ can be analytically continued to a function $\Psi(\tau, \overline{\tau}')$ with τ' at least slightly away from τ. It does not seem to be a standard fact[13] that Ψ analytically continues to a holomorphic function on $\mathcal{T} \times \mathcal{T}$ (with invariance only under one diagonal copy of the mapping class group). However, this is true in genus 1, since the partition function can be defined as $\operatorname{Tr} q^{L_0} \overline{q}'^{\overline{L}_0}$, where we can take q and q' to be independent complex variables of modulus less than 1. It seems very plausible that the statement is actually true for all values of the genus, since one can move on Teichmuller space by "cutting" on a circle and inserting $q^{L_0} \overline{q}'^{\overline{L}_0}$. If so, the partition function of the CFT can always be interpreted as a vector[14] in the Chern-Simons Hilbert space $\mathcal{H}_C(k_L, k_R)$.

If we are given a theory of three-dimensional gravity, possibly coupled to other fields, the partition function of the dual CFT is a wavefunction $\Psi(\tau, \overline{\tau}')$ which, according to the conjecture just stated, is a vector in $\mathcal{H}_C(k_L, k_R)$. Any gravitational theory of the same central charges leads to another vector in the same space.

From this point of view, it seems that we should not claim, as was done in [6], that $\mathcal{H}_C(k_L, k_R)$ is a space of physical states that are physically meaningful in pure three-dimensional gravity. Rather, a particular bulk gravitational theory, such as pure gravity, gives rise to a particular dual CFT whose partition function gives a definite vector in $\mathcal{H}_C(k_L, k_R)$. Another gravitational theory, perhaps with matter fields, whose dual CFT has the same values of the central charges, will lead to a dual partition function that is another vector in the same space. Thus, $\mathcal{H}_C(k_L, k_R)$ is in a sense a universal target for gravitational theories – with arbitrary matter fields – of given central charges.

We have formulated this for a particular Riemann surface C, but in either the gravitational theory or the dual CFT, C can vary and there is a nice behavior when C degenerates. So it is more natural to think of this as a structure that is defined for all Riemann surfaces. In conformal theory, this perspective is described in [41].

---

[13] However, G. Segal has obtained results in this direction.

[14] It may be necessary here to extend $\mathcal{H}_C(k_L, k_R)$ to a space of forms on $\mathcal{T} \times \mathcal{T}$ that are invariant under the action of the mapping class group but not necessarily square-integrable.

2.5. Analog For Supergravity

Here we will discuss the extension of some of these ideas to three-dimensional supergravity.

We consider primarily minimal supergravity, corresponding to the case that the boundary CFT has $\mathcal{N} = 1$ supersymmetry for left-movers or right-movers or perhaps both. Thus, for, say, the left-movers, the Virasoro algebra is replace by an $\mathcal{N} = 1$ super-Virasoro algebra. The symmetry algebra generated by $L_{\pm 1}$ and L_0 is extended (in the Neveu-Schwarz sector) to a superalgebra that also includes the fermionic generators $\mathcal{G}_{\pm 1/2}$. This is the Lie superalgebra of the supergroup $OSp(1|2)$, whose bosonic part is $Sp(2,\mathbb{R})$, or equivalently $SL(2,\mathbb{R})$. In particular, since the operators $\mathcal{G}_{\pm 1/2}$ transform in the two-dimensional representation of $SL(2,\mathbb{R})$, the relevant group is definitely $SL(2,\mathbb{R})$ (or possibly a covering of it), not its quotient $SO(2,1)$.

For definiteness, consider a two-dimensional CFT with $(0,1)$ supersymmetry, that is, with $\mathcal{N} = 1$ supersymmetry for right-movers and none for left-movers. Then left-movers have an ordinary Virasoro symmetry and right-movers have an $\mathcal{N} = 1$ super-Virasoro symmetry. Such a theory can be dual to a three-dimensional supergravity theory, which classically can be described by a Chern-Simons gauge theory in which the gauge supergroup is $SO(2,1) \times OSp(1|2)$, or possibly a cover thereof. For brevity, we will here assume that the gauge group is precisely $SO(2,1) \times OSp(1|2)$. The action is the obvious analog of (2.15):

$$\begin{aligned} I =& k_L I_L + k_R I_R \\ =& \frac{k_L}{4\pi}\int \mathrm{tr}\left(A_L \wedge dA_L + \frac{2}{3} A_L \wedge A_L \wedge A_L\right) - \frac{k_R}{4\pi}\int \mathrm{str}\left(A_R \wedge dA_R + \frac{2}{3} A_R \wedge A_R \wedge A_R\right). \end{aligned} \tag{2.28}$$

Here A_L is an $SO(2,1)$ gauge field, A_R is an $OSp(1|2)$ gauge field, and str is the supertrace in the adjoint representation of $OSp(1|2)$.

We want to generalize the analysis of section 2.1 to determine the allowed values of k_L and k_R. There is actually almost nothing to do. A_L is simply an $SO(2,1)$ gauge field, so k_L must be an integer. As for A_R, we can for topological purposes replace the supergroup $OSp(1|2)$ by its bosonic reduction $SL(2,\mathbb{R})$, since the fermionic directions are infinitesimal and carry no topology. So we can borrow the result of (2.8):

$$\begin{aligned} k_L &\in \mathbb{Z} \\ k_R &\in \frac{1}{4}\mathbb{Z}. \end{aligned} \tag{2.29}$$

We still have $(c_L, c_R) = (24k_L, 24k_R)$, since the Brown-Henneaux computation of the central charge depends only on the bosonic part of the action. So c_L must be a multiple of 24, as before, but now it seems that c_R should be a multiple of 6.

This is not the result that one might hope for, because holomorphic factorization in $\mathcal{N} = 1$ superconformal field theories requires that c_R should be a multiple of 12, not 6. So half of the seemingly allowed values of c_R are difficult to interpret in the spirit of this paper. Replacing $SO(2,1) \times OSp(1|2)$ by a covering group would only make things worse, as in the bosonic case.

There are many conceivable ways to interpret this result, including the possibility that our assumptions have been too optimistic. However, one additional possibility seems worthy of mention here. Part of the structure of a superconformal field theory is that there are Ramond-sector vertex operators. They introduce a "twist" in the supercurrent, which has a monodromy of -1 around a point at which a Ramond vertex operator is inserted. Let us assume that the superconformal dual of three-dimensional supergravity should be a theory in which Ramond vertex operators make sense. What is the gravitational dual of an insertion of such a vertex operator?

As in fig. 2, points in the conformal boundary of spacetime at which a Ramond vertex operator is inserted must be connected in the bulk by lines – such as the line labeled L in the figure – around which the fermionic fields of $OSp(1|2)$ (that is, the gravitinos) have a monodromy -1. It is plausible to interpret these lines, which we will call Ramond lines, as world-lines of Ramond-sector black holes. A spacetime history containing a black hole trajectory is of course more complicated than a spacetime with a simple line drawn in it. But for our present purposes, which are purely topological, the difference may not be important. There can also be Neveu-Schwarz sector black holes (and black holes in purely bosonic gravity), but we are about to discuss an effect which seems special to Ramond-sector black holes.

So now we have a new problem. We want to define

$$I_R = \frac{1}{4\pi} \int_W \mathrm{str} \left(A_R \wedge dA_R + \frac{2}{3} A_R \wedge A_R \wedge A_R \right) \tag{2.30}$$

in the presence of a Ramond world-line L on W. In the absence of the Ramond line, we know that I_R in $OSp(1|2)$ gauge theory is defined modulo $4 \cdot 2\pi$, which is why as stated in (2.29), k_R can be a multiple of $1/4$. What happens in the presence of the Ramond line?

Rather than solving a new topological problem, we can take the following shortcut. Let W' be a new three-manifold obtained by taking a double cover of W branched over the

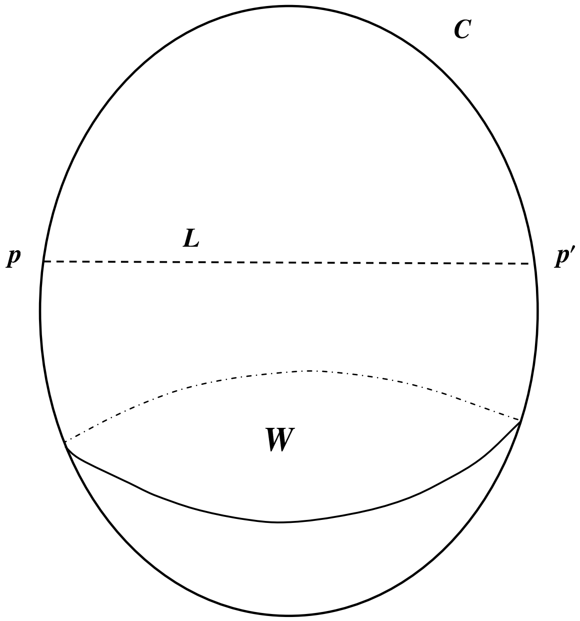


Fig. 2:

C is a Riemann surface with a pair of insertions of Ramond-sector vertex operators at points labeled p and p'. They are connected by a line L that runs in a three-manifold W whose boundary is C. Supergravity on W contributes to conformal field theory on C. The fermion fields of $OSp(1|2)$ receive a minus sign in monodromy around L.

line L. We let $I_R(W)$ be the action (2.30) and $I_R(W')$ be the corresponding action for the gauge field A_R pulled back to W'. When pulled back to W', the singularity of A_R along the Ramond line disappears, so $I_R(W')$ is defined modulo $4 \cdot 2\pi$. There is no better way to define $I_R(W)$ in the presence of a Ramond line than to say that $I_R(W) = I_R(W')/2$. So $I_R(W)$ is defined modulo $2 \cdot 2\pi$.

This means that k_R should be a multiple of 1/2, not 1/4. So in other words, if including Ramond lines is the right thing to do, we get

$$\begin{aligned} k_L &\in \mathbb{Z} \\ k_R &\in \frac{1}{2}\mathbb{Z}, \end{aligned} \tag{2.31}$$

and hence c_L and c_R are multiples of 24 and 12, respectively.

Unfortunately, the above "derivation" is little more than a scenario to try to justify the answer that we hoped for. However, a good pragmatic reason to focus on the case that c_R is a multiple of 12 is that there are interesting candidate superconformal field theories (SCFT's) in that case, as we discuss in section 3. There are no obvious interesting candidates at $c_R = 6, 18$, etc.

In the supersymmetric case it is convenient to express the Chern-Simons coupling k as $k = k^*/2$, where we will focus on the case that k^* is an integer. In terms of k^*, the central charge is $c = 12k^*$.

Some Generalizations

We conclude this section by briefly mentioning some simple generalizations.

First of all, $(1,1)$ supersymmetry in two dimensions, with $\mathcal{N} = 1$ super-Virasoro symmetry for both left-movers and right-movers, is dual to three-dimensional supergravity theories related to $OSp(1|2) \times OSp(1|2)$ Chern-Simons gauge theory. If one wants the left- or right-movers to have more than $\mathcal{N} = 1$ supersymmetry, one simply replaces $OSp(1|2)$ by an appropriate supergroup with more fermionic generators. For example, $OSp(2|2)$ is related to $\mathcal{N} = 2$ supersymmetry, $PSU(2|2)$ is related to what is usually called $\mathcal{N} = 4$ supersymmetry (with the "small" $\mathcal{N} = 4$ superconformal algebra), and $OSp(r|2)$ with $r > 2$ is related to theories with "large" $\mathcal{N} = k$ superconformal algebras. (The most widely studied case of these algebras is $r = 4$; for example, see [42].)

It is also possible to consider other extended chiral algebras, apart from superconformal algebras. For example, one can start in three dimensions with an $SL(3, \mathbb{R})$ Chern-Simons gauge theory, which plausibly may be related to some sort of three-dimensional W_3 gravity theory in much the same (not fully understood) way that $SL(2, \mathbb{R})$ Chern-Simons theory is related to ordinary three-dimensional gravity. A dual CFT would very likely then have a W_3 chiral algebra. An analogous statement plausibly holds for many extended chiral algebras.

3. Partition Functions

In this section, we will determine what we propose to be the exact spectrum of physical states of three-dimensional gravity or supergravity with negative cosmological constant, in a spacetime asymptotic at infinity to Anti de Sitter space. Equivalently, we will determine the genus one partition function of the dual CFT.

In all cases, we work at the values of ℓ/G at which holomorphic factorization is possible. These values were related to gauge theory in section 2. We assume holomorphic factorization, and mainly consider only the holomorphic sector of the theory. The full partition function is the product of the function we determine and a similar antiholomorphic function. These partition functions have been studied before [31,32], with different motivation.

3.1. The Bosonic Case

We begin with the bosonic case. What are the physical states of pure gravity in a spacetime asymptotic at infinity to AdS_3?

Since there are no gravitational waves in the theory, the only state that is obvious at first sight is the vacuum, corresponding in the classical limit to Anti de Sitter space. In a conformal field theory with central charge $c = 24k$, the ground state energy is $L_0 = -c/24 = -k$. The contribution of the ground state $|\Omega\rangle$ to the partition function $Z(q) = \operatorname{Tr} q^{L_0}$ is therefore q^{-k}.

Of course, there is more to the theory than just the ground state. According to Brown and Henneaux [21], a proper treatment of the behavior at infinity leads to the construction of a Virasoro algebra that acts on the physical Hilbert space. The Virasoro generators L_n, $n \geq -1$ annihilate $|\Omega\rangle$, but by acting with $L_{-2}, L_{-3}, \ldots$, we can make new states of the general form $\prod_{n=2}^{\infty} L_{-n}^{s_n}|\Omega\rangle$, with energy $-k+\sum_n ns_n$. (We assume that all but finitely many of the s_n vanish.) If these are the only states to consider, then the partition function would be

$$Z_0(q) = q^{-k}\prod_{n=2}^{\infty}\frac{1}{1-q^n}. \tag{3.1}$$

This cannot be the complete answer, because the function $Z_0(q)$ is not modular-invariant. There must be additional states such that $Z_0(q)$ is completed to a modular-invariant function.

Additional states are expected, because the theory also has BTZ black holes. The main reason for writing the present paper, after all, is to understand the role of the BTZ black holes in the quantum theory. We will assume that black holes account for the difference between the naive partition function $Z_0(q)$ and the exact one $Z(q)$. To use this assumption to determine $Z(q)$, we need to know something about the black holes.

The classical BTZ black hole is characterized by its mass M and angular momentum J. In terms of the Virasoro generators,

$$\begin{aligned} M =& \frac{1}{\ell}(L_0 + \overline{L}_0) \\ J =& (L_0 - \overline{L}_0), \end{aligned} \tag{3.2}$$

so $L_0 = (\ell M + J)/2$, $\overline{L}_0 = (\ell M - J)/2$. The classical BTZ black hole obeys $M\ell \geq |J|$, or $L_0, \overline{L}_0 \geq 0$. The BTZ black hole is usually studied in the absence of the gravitational Chern-Simons coupling, that is for $k_L = k_R = k$. Its entropy is $S = \pi(\ell/2G)^{1/2}\left(\sqrt{M\ell - J} + \sqrt{M\ell + J}\right)$. (This entropy was first expressed in terms of two-dimensional conformal field theory in [13].) With $\ell/G = 16k$ as in (2.19), this is equivalent to $S = 4\pi\sqrt{k}\left(\sqrt{L_0} + \sqrt{\overline{L}_0}\right)$. For the holomorphic sector, the entropy is therefore

$$S_L = 4\pi\sqrt{k_L L_0}, \tag{3.3}$$

and similarly for the antiholomorphic sector.

There is no classical BTZ black hole with $L_0 < 0$, and the entropy of such a black hole is zero if $L_0 = 0$. We will take this as a suggestion that quantum states corresponding to black holes exist only if $L_0 > 0$, that is $L_0 \geq 1$. This means that the exact partition function $Z(q)$ should differ from the function $Z_0(q)$ in (3.1) by terms of order q:

$$Z(q) = q^{-k}\prod_{n=2}^{\infty}\frac{1}{1-q^n} + \mathcal{O}(q). \tag{3.4}$$

Admittedly, we are here trying to squeeze more information from the classical result than is justified. But it turns out that a modular-invariant partition function of this form exists and is unique. This result (which is due to Höhn [31]) follows from the fact that the moduli space $\mathcal{M}_1$ of Riemann surfaces of genus 1 is itself a Riemann surface of genus 0, in fact parametrized by the j-function. If E_4 and E_6 are the usual Eisenstein series of weights 4 and 6, then $j = 1728E_4^3/(E_4^3 - E_6^2)$. Its expansion in powers of q is

$$j(q) = q^{-1} + 744 + 196884q + 21493760q^2 + 864299970q^3 + 20245856256q^4 + \dots. \tag{3.5}$$

Actually, it is more convenient to use the function

$$J(q) = j(q) - 744 = q^{-1} + 196884q + 21493760q^2 + \dots, \tag{3.6}$$

which likewise parametrizes the moduli space.

The J-function has a pole at $q = 0$ and no other poles. The statement that J parametrizes the moduli space means precisely that any modular-invariant function can be written as a function of J. The partition function $Z(q)$ has a pole at $q = 0$, that is at $J = \infty$. This pole is of order k, and the partition function has no other poles. Any holomorphic function of J that has no singularity except at $J = \infty$ is a polynomial in J.

In the present case, as the pole in $Z(q)$ at $q=0$ is of order k, Z must be a polynomial in J of degree k. Thus

$$Z(q)=\sum_{r=0}^{k} f_r J^r, \tag{3.7}$$

with some coefficients f_r. These $k+1$ coefficients can be adjusted in a unique fashion to ensure that $Z(q)$ takes the form (3.4), or in other words to ensure that the terms in $Z(q)$ of order q^{-n}, $n=0,\dots,k$, coincide with the naive function $Z_0(q)$.

When we do this, we get a function that we will call $Z_k(q)$, $k=1,2,3,\dots$. This function is our candidate for the generating function that counts the quantum states of three-dimensional gravity in a spacetime asymptotic to AdS_3. For example, for $k=1$ we have simply $Z_1(q)=J(q)$, and the next few examples are

$$\begin{aligned}
Z_2(q) =& J(q)^2-393767\\
=& q^{-2}+1+42987520q+40491909396q^2+\dots\\
Z_3(q) =& J(q)^3-590651J(q)-64481279\\
=& q^{-3}+q^{-1}+1+2593096794q+12756091394048q^2+\dots\\
Z_4(q) =& J(q)^4-787535J(q)^2-8597555039J(q)-644481279\\
=& q^{-4}+q^{-2}+q^{-1}+2+81026609428q+1604671292452452276q^2+\dots.
\end{aligned} \tag{3.8}$$

Following [31], we refer to a holomorphic CFT with $c=24k$ and partition function $Z_k(q)$ as an extremal CFT. According to our proposals, the dual of three-dimensional gravity should be an extremal CFT.

As was already noted in the introduction, Frenkel, Lepowsky, and Meurman constructed [24,25] an extremal CFT with $k=1$, that is, a holomorphic CFT with $c=24$ and partition function $J(q)=Z_1(q)$. They also conjectured its uniqueness. If that conjecture as well as the ideas in the present paper are correct, then the FLM theory must be the dual to quantum gravity for $k=1$. Unfortunately, as also noted in the introduction, for $k>1$, extremal CFT's are not known, though their possible existence has been discussed in the literature [31-33] for reasons not related to three-dimensional gravity. Our reasoning in this paper suggests that such theories should exist and be unique for each k.

The main point of the FLM construction was that their theory has as a group of symmetries the Fischer-Griess monster group $\mathbb{M}$, the largest of the sporadic finite groups. Arguably, the FLM theory is the most natural known structure with $\mathbb{M}$ symmetry. The coefficients in the q-expansion of the J-function are integers for number-theoretic reasons, but the FLM construction gave a new perspective on why they are positive; this is a property of the partition function of any CFT. It also gave a new perspective on why these coefficients are so large; this follows from $\mathbb{M}$ symmetry, since $\mathbb{M}$ does not have small representations.

Indeed, the original clue to the FLM construction was the observation by J. McKay that the first non-trivial coefficient 196884 of the function $J(q)$ is nearly equal to the smallest dimension 196883 of a non-trivial representation of $\mathbb{M}$. (This observation was later greatly generalized [43,44].) The FLM interpretation is that 196884 is the number of operators of dimension 2 in their theory. One of these operators is the stress tensor, while the other 196883 are primary fields transforming in the smallest non-trivial representation of $\mathbb{M}$.

In our interpretation, the 196883 primaries are operators that (when combined with suitable anti-holomorphic factors) create black holes. It is illuminating to compare the number 196883 to the Bekenstein-Hawking formula. An exact quantum degeneracy of 196883 corresponds to an entropy of $\ln 196883 \cong 12.19$. By contrast, the Bekenstein-Hawking entropy at $k = 1$ and $L_0 = 1$ is $4\pi \cong 12.57$. We should not expect perfect agreement, because the Bekenstein-Hawking formula is derived in a semiclassical approximation which is valid for large k.

Agreement improves rapidly if one increases k. For example, at $k = 4$, and again taking $L_0 = 1$, the exact quantum degeneracy of primary states is 81026609426, according to eqn. (3.8). (Two of the states at this level are descendants.) This corresponds to an entropy $\ln 81026609426 \cong 25.12$, compared to the Bekenstein-Hawking entropy $8\pi \cong 25.13$. Shortly we will compute the entropy in the large k limit.

Our interpretation is that a primary state $|\Lambda\rangle$ represents a black hole, while a descendant $\prod_{n=1}^{\infty} L_{-n}^{s_n}|\Lambda\rangle$ describes a black hole embellished by boundary excitations. These boundary excitations are the closest we can come in $2+1$ dimensions to the gravitational waves that a black hole can interact with in a larger number of dimensions. If this interpretation is correct, then for an exact count of black hole states, we should count primaries only. However, as the above examples illustrate, in practice this issue has only a very slight effect on the black hole degeneracies. The boundary excitations contribute to the entropy an amount that is independent of k and thus negligible in the regime where we compare to the Bekenstein-Hawking entropy, and also, numerically, negligible even for small k. The separation between the black hole and the boundary excitations is best-motivated for large k.

Alternative Formula

Now we will present an alternative formula for the partition functions $Z_k(q)$. This formula avoids the large coefficients involved in writing the Z_k as a polynomial in J, and will make some properties of the Z_k more manifest.

Let $q = \exp(2\pi i\tau)$ with τ in the upper half plane. If p is a prime number, then the Hecke operators acting on functions $F(\tau)$ are defined essentially by

$$T'_pF(\tau) = F(p\tau) + \sum_{b=0}^{p-1} F((\tau+b)/p). \tag{3.9}$$

(What we call T'_p is p times the Hecke operator T_p as usually defined. See [45] for an introduction to Hecke operators and [29,30] for their use in the present subject.) More generally, for any positive integer t, we define

$$T'_tF(\tau) = \sum_{d|t}\sum_{b=0}^{d-1} F((t\tau+bd)/d^2). \tag{3.10}$$

For $t = 0$, we define $T'_0F(\tau) = 1$. If $F(\tau)$ is modular-invariant, then so is $T'_tF(\tau)$. The definition of Hecke operators generalizes naturally to modular forms, but we will not need this.

An immediate consequence of the definition is that, for any $F(\tau)$ that is invariant under $\tau \to \tau + 1$, and so has a Laurent expansion in powers of $q = \exp(2\pi i\tau)$, the coefficients in the q-expansion of $T'_nF(\tau)$ are linear combinations with positive integer coefficients of the q-expansion coefficients of $F(\tau)$. The FLM construction shows that the q-expansion coefficients of the J-function are non-negative integers and the coefficients of positive powers of q are dimensions of non-trivial (reducible) monster representations. Hence the same is true of T'_nJ.

An important special case of how the Hecke operators act on q-expansion coefficients is that if $F(\tau) = q^{-1} + \mathcal{O}(q)$, then

$$T'_nF(\tau) = q^{-n} + \mathcal{O}(q). \tag{3.11}$$

Now we can give an alternative description of the partition functions $Z_k(\tau)$. We recall that these functions are uniquely determined by being modular-invariant and agreeing up to order q with the function $Z_0 = q^{-k}\prod_{n=2}^{\infty}(1-q^n)^{-1}$. To find a function with these properties, we simply expand

$$Z_0 = \sum_{r=-k}^{\infty} a_rq^r, \tag{3.12}$$

and then we let

$$Z_k(\tau) = \sum_{r=0}^{k} a_{-r} T'_r J(\tau). \tag{3.13}$$

For example,

$$Z_2(\tau) = (T'_2 + T'_0)J(\tau) = J(2\tau) + J(\tau/2) + J((\tau+1)/2) + 1. \tag{3.14}$$

This representation makes it clear that, just like the $T'_r J$, the functions Z_k share some properties with the J-function: the q-expansion coefficients are non-negative, and the coefficients of positive powers of q are dimensions of non-trivial monster representations.

This makes it tempting to speculate that the CFT's dual to three-dimensional gravity have $\mathbb{M}$ symmetry for all k. If so, the $\mathbb{M}$ symmetry is invisible in classical General Relativity and acts on the microstates of black holes. (However, the analogous conjecture for supergravity, which would involve the Conway group, appears to be untrue, as we will see in section 3.3.)

As an illustration of the usefulness of the expression for the partition function in terms of Hecke operators, we will use it to compare with the Bekenstein-Hawking entropy. We consider the semiclassical limit $k, L_0 \to \infty$, with $r = L_0/k$ positive and fixed. We take $r = p/q$ to be a rational number and assume that k is divisible by q, so that L_0 is an integer n. If we write

$$J(q) = \sum_{m=-1}^{\infty} c_m q^m, \tag{3.15}$$

then a formula of Petersson and Rademacher gives the asymptotic behavior

$$\ln c_m \sim 4\pi\sqrt{m} - \frac{3}{4}\ln m - \frac{1}{2}\ln 2 + \dots. \tag{3.16}$$

Let

$$Z_k(\tau) = \sum_{n=-k}^{\infty} b_{k,n} q^n. \tag{3.17}$$

We want to determine the behavior of $b_{k,n}$ for large k and n with $r = n/k$ fixed. To evaluate the partition function using the formula (3.13), let us first look at the contribution from $r = k$, that is $T'_k J(\tau)$. In evaluating $T'_k J$ from the definition (3.10), the dominant term (for large k and n) is the term with $d = k$. This contribution to $b_{k,n}$ is $b^0_{k,n} \sim k c_{kn}$, and hence, using the Petersson-Rademacher formula,

$$\ln b^0_{k,n} \sim 4\pi\sqrt{kn} + \frac{1}{4}\ln k - \frac{3}{4}\ln n - \frac{1}{2}\ln 2 + \dots. \tag{3.18}$$

The first term is the Bekenstein-Hawking entropy (3.3). The additional terms in (3.18), as well as the remaining contributions in (3.13) with $r < k$ that we have omitted in deriving (3.18), do not modify the Bekenstein-Hawking formula in the limit of large k, fixed n/k, but give interesting subleading corrections that will be studied elsewhere [46].

Another method to get similar results (expressing black hole degeneracies in terms of coefficients of the singular part of the partition function) is the Farey tail expansion [34]. See also [47].

The FLM Construction

Consider a holomorphic CFT with $c = 24$. If the number of primary fields of dimension 1 is s, then the expansion of the partition function near $q = 0$ begins $Z(q) = q^{-1} + s + \mathcal{O}(q)$, and modular invariance implies that the genus 1 partition function is precisely $Z(q) = J(q) + s = q^{-1} + s + 196884q + \ldots$. Of the 196884 fields of dimension 2, one is a descendant of the identity and s are descendants of dimension 1 primaries. Hence there are $196883 - s$ primary fields of dimension 2, too few to furnish a non-trivial representation of the monster group unless $s = 0$.

For this reason, Frenkel, Lepowsky, and Meurman [24] aimed to construct a holomorphic CFT of $c = 24$ with no primary fields of dimension 1, hoping that it would have monster symmetry. Their construction was made as follows. The Leech lattice is an even, unimodular lattice of rank 24 with no vector of length squared less than 4. One can construct a holomorphic theory of $c = 24$ by compactifying 24 chiral bosons X_i, $i = 1, \ldots, 24$ using any even unimodular lattice. If one uses the Leech lattice, then because it contains no vector of length squared less then 4, a primary field of the form $\exp(ip \cdot X)$ has dimension at least 2, as desired. However, this lattice theory has 24 primary fields ∂X_i of dimension 1. Aiming to eliminate those fields, FLM considered an orbifold[15] by the $\mathbb{Z}_2$ symmetry $X_i \to -X_i$ which eliminates the dimension 1 primaries. As it turns out, this particular orbifold preserves modular invariance, does not add dimension 1 primaries in twisted sectors, and gives the hoped-for theory with monster symmetry.

FLM conjectured that the model they constructed is the unique holomorphic CFT with partition function J. If this conjecture is correct, it implies that many other constructions give the same theory. There are [49] many orbifolds of the Leech lattice theory that give holomorphic CFT's with no primary of dimension 1. According to the FLM uniqueness conjecture, all of these theories are isomorphic to their monster theory. Moreover, the Leech lattice theory itself can be obtained as an orbifold theory starting from another even unimodular lattice of rank 24, so other lattices can be used as starting points as well.

---

[15] Their work preceded the general study of orbifolds in string theory [48], though some related stringy constructions, such as the Ramond and Neveu-Schwarz sectors of string theory, were already known.

Aiming to imitate the FLM construction for $k = 2$, one might start with an even unimodular rank 48 lattice with no vector of length squared less than 6. Three such lattices are known; two are described in [50] and a third has been constructed more recently [51]. (There is no reason to believe that these are the only three rank 48 lattices with this property; there may be a vast number of them.) In the theory of 48 free chiral bosons X_i compactified using such a lattice, a primary field $\exp(ip \cdot X)$ has dimension at least 3, but there are dimension 1 primaries ∂X_i and dimension 2 primaries $\partial X_i \partial X_j - \frac{1}{48}\delta_{ij}\partial X \cdot \partial X$. A simple $\mathbb{Z}_2$ orbifold eliminates the dimension 1 primaries, as in the FLM case, but not the dimension 2 primaries. One may attempt to find a more complicated orbifolding construction to remove the unwanted primaries. Though all three lattices have interesting discrete symmetry groups, it appears that there is no anomaly-free subgroup that removes all dimension 2 primaries.[16] Many other orbifolds can be considered, such as the symmetric product of k copies of the $k = 1$ monster theory, but it appears difficult to remove all primaries of low dimension.

Optimistically speaking, this situation might be compared to current algebra of a simply-laced compact Lie group G. At level 1, one has the Frenkel-Kac-Segal construction of current algebra of G via free bosons. At higher integer level, the theory still exists, but generically has no equally straightforward realization in terms of free fields. Perhaps the situation is somewhat similar for extremal CFT's.

Further Remarks

We have emphasized here the partition function, because it is what we can determine for general k. However, if one can actually describe the relevant CFT – as conjecturally we can at $k = 1$ via the FLM construction – this gives much more than a partition function. In this case, by computing matrix elements of primary fields, we get a detailed description of the black hole quantum mechanics.

Above 2+1 dimensions, a black hole can form out of radiation and ordinary matter. In pure gravity in 2+1 dimensions, the only thing that a black hole can form from is, roughly speaking, smaller black holes. To be more precise, let Φ_i, $i = 1, \ldots, w$ be primary fields of dimension not much greater than k that individually could, in acting on the vacuum, create black holes of fairly small mass. A product of such fields acting on the vacuum at prescribed points

---

[16] I am grateful for assistance from G. Nebe in investigating this question.

$$\prod_{i=1}^{w} \Phi_i(z_i)|\Psi\rangle \tag{3.19}$$

may create a black hole of large mass, that is, a primary state of large energy or a descendant of such a state. Evaluating such matrix elements is the closest analog we can find, in the present model, to describing the formation of a large mass black hole from matter. Unfortunately, at the moment, we can perform such computations only for $k = 1$, since for other values of k we do not know the CFT. Ideally, one would like to describe the CFT for all k and investigate the semiclassical limit of large k.

If our hypotheses are correct, the following may be the most significant difference between three-dimensional pure gravity and a more realistic theory of black holes in $3 + 1$ dimensions. In the present model, as the holomorphic fields are all conserved currents, and the energy levels are integers, the dynamics is integrable and periodic on a short time scale. This is certainly not expected for black holes in general. An embedding of three-dimensional gravity in a larger system, such as a string theory, would be expected to give a non-integrable deformation of the model.

A Conundrum

We have here considered the holomorphic sector of a CFT, but the CFT dual of three-dimensional gravity has both holomorphic and antiholomorphic degrees of freedom. Let us therefore discuss a puzzle (stressed by J. Maldacena) that arises when one combines the holomorphic and antiholomorphic degrees of freedom. We let $|\Omega\rangle$, $|\widetilde{\Omega}\rangle$ be the ground states in the holomorphic and antiholomorphic sectors, and let $|\Phi\rangle$ and $|\widetilde{\Phi}\rangle$ denote primary states other than the ground state. The state $|\Omega\rangle \otimes |\widetilde{\Omega}\rangle$ and its descendants correspond, according to our picture, to Anti de Sitter space and its boundary excitations. A state $|\Phi\rangle \otimes |\widetilde{\Phi}\rangle$, or a descendant thereof, corresponds to a BTZ black hole, perhaps with boundary excitations. But what do we make of states $|\Omega\rangle \otimes |\widetilde{\Phi}\rangle$ or $|\Phi\rangle \otimes |\widetilde{\Omega}\rangle$, and their descendants? Such states are trying to be Anti de Sitter space for holomorphic variables and black holes for antiholomorphic variables, or vice-versa.

In classical three-dimensional gravity, there is no satisfactory solution that has this interpretation. We can, however, see what is involved in trying to make one. First of all, three-dimensional Anti de Sitter space AdS$_3$ is simply the universal cover of the $SL(2, \mathbb{R})$ group manifold, with the symmetry $SL(2, \mathbb{R}) \times SL(2, \mathbb{R})$ (or rather a cover thereof) coming from the left and right action of $SL(2, \mathbb{R})$ on itself. The conformal boundary of the universal cover of AdS$_3$ is a cylinder $\mathbb{R} \times S^1$. We can parametrize it by $-\infty < \tau < \infty$, $0 \leq \sigma \leq 2\pi$, with conformal structure described by the metric $ds^2 = d\tau^2 - d\sigma^2$. The BTZ black hole is the quotient of (a cover of) $SL(2, \mathbb{R})$ by a subgroup $\mathbb{Z} \subset SL(2, \mathbb{R}) \times SL(2, \mathbb{R})$ that is not simply a subgroup of one factor or the other. This quotient has conformal boundary that is again a cylinder, conformally equivalent to the boundary of Anti de Sitter space itself. That is why the BTZ black hole can be regarded as an excitation of Anti de Sitter space.

To describe a state $|\Omega\rangle \otimes |\Phi\rangle$ (or its parity conjugate) at the classical level, we want a solution of Einstein's equations that is invariant under precisely one of the two factors of the $SL(2, \mathbb{R}) \times SL(2, \mathbb{R})$ symmetry. This will reflect the fact that $|\Omega\rangle$ is $SL(2, \mathbb{R})$-invariant and $|\Phi\rangle$ is not. A space with just one of the $SL(2, \mathbb{R})$ symmetries is AdS$_3/\mathbb{Z}$, where $\mathbb{Z}$ is a subgroup of one $SL(2, \mathbb{R})$ factor. The quotient AdS$_3/\mathbb{Z}$ is a perfectly good solution of Einstein's equations in bulk, but its asymptotic behavior at infinity does not quite agree with that of AdS$_3$. The boundary at infinity is a cylinder, but the compact direction in the cylinder is lightlike rather than spacelike. For our approach to three-dimensional gravity in the present paper to be correct, something like this quotient AdS$_3/\mathbb{Z}$ has to make sense at the quantum level, perhaps because of a small quantum correction that makes the compact direction in the cylinder effectively spacelike. Note that the metric $ds_\epsilon^2 = d\tau^2 - \epsilon d\sigma^2$ induces on the cylinder a conformal structure that is independent of ϵ up to isomorphism as long as $\epsilon > 0$.

What we have described is an analog for black holes of a statement made in section 2.4: to ensure holomorphic factorization, the sum over topologies probably must be extended to include configurations that are difficult to interpret classically. In the present case, the analog of the sum over topologies is the sum over states with or without the presence of a black hole.

3.2. The Supersymmetric Case

Now we consider the analog for supergravity. The analysis is a little more complicated and in some ways the results are less satisfactory.

A rather similar problem has been treated by Höhn [31]. He considered not a superconformal field theory, but a more general holomorphic theory with bosonic operators of integral dimension and fermionic operators of half-integral dimension. In that case, the partition function can have a pole at the Ramond cusp, and is given by a Laurent series in $j_\theta = K - 24$ rather than a polynomial.

We continue to assume holomorphic factorization, and we assume that the holomorphic part of the boundary CFT is an $\mathcal{N} = 1$ superconformal field theory (SCFT). This means (in whatever not well understood sense gravity is related to gauge theory) that the gauge group of the bulk theory has a factor $OSp(1|2)$. From a classical point of view, the $OSp(1|2)$ bundle over a three-manifold W endows W with a spin structure. In the AdS/CFT correspondence, W has for conformal boundary a Riemann surface C, and the spin structure on W determines one on C. One is instructed in the AdS/CFT correspondence to specify C, including its spin structure, and sum over all choices of W. (Depending on the theory, the antiholomorphic degrees of freedom may themselves be supersymmetric and endowed with a choice of spin structure.) We will assume that it is physically appropriate to specify the spin structure on C rather than summing over it (just as one specifies the complex structure of C), though at the end of section 3.3, we discuss what happens if one sums over spin structures.

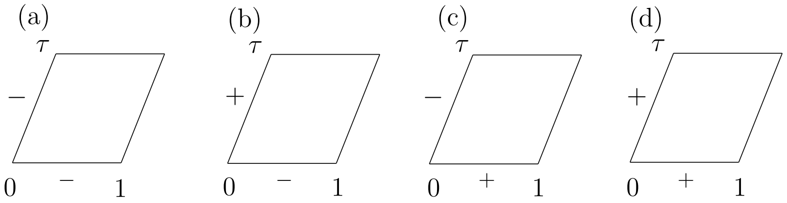


Fig. 3:

We represent a genus 1 Riemann surface C as the quotient of the complex plane by the lattice generated by complex numbers 1 and τ, where τ lies in the upper half plane $\mathcal{H}$. There are four possible spin structures, classified by whether the fermions are periodic or antiperiodic around the "horizontal" and "vertical" cycles. This is indicated by labeling the horizontal and vertical directions with a + sign for periodicity or a − sign for antiperiodicity.

We focus on the case that C has genus 1, in which case there are four possible spin structures (fig. 3). The path integrals for these four spin structures can be interpreted in terms of traces in two Hilbert spaces, known as the Neveu-Schwarz (NS) and Ramond (R) Hilbert spaces $\mathcal{H}_{\mathrm{NS}}$ and $\mathcal{H}_{\mathrm{R}}$, respectively.

Traces in $\mathcal{H}_{\mathrm{NS}}$ are constructed using the spin structures, shown in (a) and (b) in the figure, in which fermions are antiperiodic in the horizontal direction. If additionally the fermions are antiperiodic in the vertical direction, as in (a), then the path integral computes

$$F(\tau) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{NS}}}\, q^{L_0}, \tag{3.20}$$

where $q = \exp(2\pi i\tau)$. We call this the NS partition function. If they are periodic in the vertical direction, as in (b), the path integral computes instead

$$G(\tau) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{NS}}}(-1)^F q^{L_0}, \tag{3.21}$$

where $(-1)^F$ is the operator that is $+1$ on bosonic states, including the ground state of the NS sector, and -1 on fermionic states.

Traces in $\mathcal{H}_{\mathrm{R}}$ are computed using spin structures, shown in (c) and (d) in the figure, in which fermions are periodic in the horizontal direction. If the fermions are antiperiodic in the vertical direction, the path integral computes an ordinary trace

$$H(\tau) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{R}}}\, q^{L_0}. \tag{3.22}$$

If fermions are periodic in both directions, the path integral gives

$$\chi = \mathrm{Tr}_{\mathcal{H}_{\mathrm{R}}}\, (-1)^F q^{L_0} = \mathrm{Tr}\, (-1)^F. \tag{3.23}$$

Both $H(\tau)$ and χ are severely constrained by the fact that in the R sector, one of the superconformal generators, which we will $\mathcal{G}_0$, commutes with L_0 and obeys $\mathcal{G}_0^2 = L_0$. This implies first of all that $L_0 \geq 0$ in the R sector, so $H(\tau)$ has no pole at $q = 0$. Further, the action of $\mathcal{G}_0$ implies that states of $L_0 > 0$ are paired between bosons and fermions. They cancel out of χ and make equal contributions to $H(\tau)$. It follows that χ is simply an integer, the trace of the operator $(-1)^F$ in the subspace with $L_0 = 0$ (or alternatively the index of the operator $\mathcal{G}_0$ mapping from bosonic states to fermionic ones). Moreover, if we expand $H(\tau)$ in powers of q

$$H(\tau) = \sum_{n=0}^{\infty} h_n q^n, \tag{3.24}$$

then all coefficients h_n are even except possibly h_0. In this expansion, only integral powers of q appear, since all fields obey integral boundary conditions in the R sector.

In the NS sector, the ground state energy is $-c/24$. In the holomorphically factorized case that we focus on, $c = 12k^*$ for an integer k^*, so the ground state energy is $-k^*/2$.

Since fermions in the NS sector are antiperiodic in the spatial direction, NS excitations may have either integer or half-integer energy above the ground state. The partition function therefore takes the general form

$$F(\tau) = q^{-k^*/2}(1 + aq^{1/2} + bq + cq^{3/2} + \ldots) = \sum_{m\in\mathbb{Z}/2,\, m\geq -k^*/2} f_m q^m. \tag{3.25}$$

In this expansion, the states in which $m + k/2$ is an integer are bosons, while the states in which $m + k/2$ is half-integral are fermions. That is so because, in the NS sector, bosonic fields are periodic and have integral excitation energies, while fermionic fields are half-integral and have half-integral excitation energies. $G(\tau)$ can therefore very explicitly be expressed in terms of the same coefficients:

$$G(\tau) = \sum_{m\in\mathbb{Z}/2,\, m\geq -k^*/2} (-1)^{2m+k^*} f_m q^m. \tag{3.26}$$

We have simply included a factor that is $+1$ on bosons and -1 on fermions. A more succinct way to say the same thing is that

$$G(\tau) = (-1)^{k^*} F(\tau + 1). \tag{3.27}$$

Thus, G is simply determined in terms of F. The same is true for H. Consider the modular transformation $\tau \to -1/\tau$, which has the effect of exchanging the horizontal and vertical directions in fig. 3 (with a reversal of orientation for one). This exchanges spin structures (b) and (c), as a result of which

$$H(\tau) = G(-1/\tau). \tag{3.28}$$

We can combine (3.27) and (3.28) to get

$$H(\tau) = (-1)^{k^*} F(-1/\tau + 1). \tag{3.29}$$

An important special case of this is

$$\lim_{\tau\to i\infty} H(\tau) = (-1)^{k^*} F(1). \tag{3.30}$$

The limit on the left hand side exists, because H has no pole at $q = 0$. The limit is just the coefficient h_0 in (3.24):

$$h_0 = (-1)^{k^*} F(1). \tag{3.31}$$

h_0 is the number of Ramond states of zero energy.

According to the above formulas, everything (except an integer χ, the supersymmetric index) can be expressed in terms of F. So it is useful to understand the modular properties of F. F is not invariant under $\tau \to \tau + 1$, as we have already seen. But, since all energies take values in $\mathbb{Z}/2$, F is invariant under $\tau \to \tau + 2$. In addition, since the modular transformation $\tau \to -1/\tau$ maps spin structure (a) itself, F is invariant under $\tau \to -1/\tau$. The two transformations $\tau \to \tau + 2$ and $\tau \to -1/\tau$ generate a subgroup of $SL(2, \mathbb{Z})$ that consists of 2×2 integral unimodular matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.32}$$

that are congruent mod 2 to one of

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.33}$$

These matrices of course act on τ in the usual fashion, $\tau \to (a\tau + b)/(c\tau + d)$. This group is sometimes called Γ_θ; it is conjugate to the group $\Gamma_0(2)$ characterized by requiring b to be even. (A third conjugate group is characterized by the condition that c should be even. Each of these three groups can be defined as the subgroup of $SL(2, \mathbb{Z})$ that leaves fixed one of the even spin structures, that is, one of the first three shown in fig. 3.)

The quotient of the upper half plane $\mathcal{H}$ by $SL(2, \mathbb{Z})$ is a curve of genus zero with one point missing. This can be seen by studying the fundamental domain (fig. 4(a)). The missing point is called the "cusp" at $\tau = i\infty$. It is because the quotient has genus zero that Riemann surfaces of genus 1 can be parametrized by a single holomorphic function, namely $J(\tau)$. As we have seen in section 3.1, this is useful because it means that the constraints of modular invariance are equivalent to the statement that the partition function is a function of J. Similarly, the quotient of the upper half plane $\mathcal{H}$ by Γ_θ is again a curve of genus zero, this time with two missing points or cusps, as explained in fig. 4(b). This means that there is a function $K(\tau)$ such that any Γ_θ-invariant function F can be expressed as a function of K.

The two cusps of $\mathcal{H}/\Gamma_\theta$ are at $\tau \to i\infty$ and $\tau = 1$, and correspond respectively to the NS sector and the R sector. For $\tau \to i\infty$, the NS partition function $F(\tau)$ has an expansion in terms of L_0 eigenvalues in the NS sector. Likewise, the Ramond partition function $H(\tau)$ can be expanded for $\tau \to i\infty$ in L_0 eigenvalues in the Ramond sector; but according to (3.29), the behavior of $H(\tau)$ for τ near $i\infty$ is the same as the behavior of $F(\tau)$ for τ near 1.

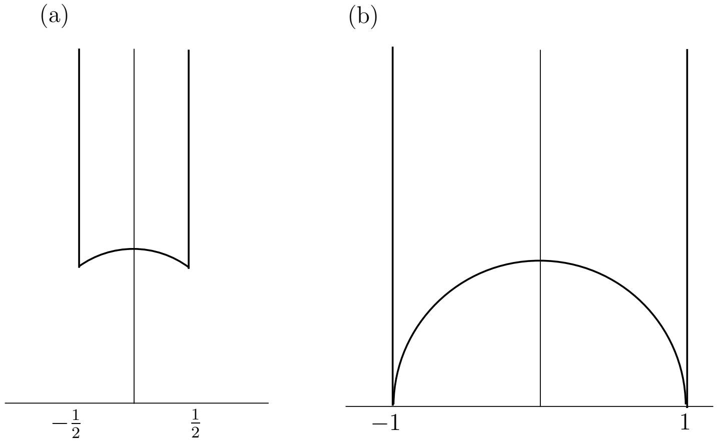


Fig. 4:

(a) A fundamental domain for the action of $SL(2,\mathbb{Z})$ on the upper half plane. By the action of $\tau \to \tau+1$, one can take $|\mathrm{Re}\,\tau| \leq 1/2$, and by the action of $\tau \to -1/\tau$, one can take $|\tau| \geq 1$. The $SL(2,\mathbb{Z})$ action identifies the left and right hand halves of the boundary of the fundamental domain, making what topologically is a Riemann surface of genus zero with one missing point at $\tau = i\infty$. This point is called the cusp. (b) The analogous fundamental domain for the action of Γ_θ. Here the symmetry $\tau \to \tau+2$ lets us reduce to $|\mathrm{Re}\,\tau| \leq 1$, and $\tau \to -1/\tau$ still lets us reduce to $|\tau| \geq 1$. The group action still identifies the left and right hand halves of the boundary of the fundamental domain, and the result is a Riemann surface of genus zero, now with two points omitted. The missing points are the Neveu-Schwarz cusp at $\tau = i\infty$ and the Ramond cusp at $\tau = \pm 1$ (those two points are equivalent under $\tau \to \tau+2$). The Ramond cusp is missing in the quotient of the upper half plane by Γ_θ because the points $\tau = \pm 1$ are not in the upper half plane, but on its boundary.

So in terms of the function F, the behavior at the two cusps is determined by the low-lying spectrum in the NS and Ramond sectors, respectively.

We therefore call these the NS and R cusps. The natural uniformizing parameter at the NS cusp is $q^{1/2} = \exp(i\pi\tau)$, where τ is the argument of the function F in (3.25). The right parameter is $\exp(i\pi\tau)$ since the symmetry is $\tau \to \tau+2$. At the R cusp, the natural parameter is $q = \exp(2\pi i\tau)$, where now τ is the argument of H.

To get a clear picture of the nature of a holomorphic function F on $\mathcal{H}/\Gamma_\theta$, it is necessary to describe the behavior near the cusps, or equivalently to describe what happens when one tries to extend F to a function on the compactification of $\mathcal{H}/\Gamma_\theta$ that is obtained by adding the cusps. This compactification is a space $\mathcal{Y} \cong \mathbb{CP}^1$.

We can explicitly describe a function K that parametrizes $\mathcal{H}/\Gamma_\theta$ and has a pole only at the NS cusp. This can be done in several ways. One formula is

$$K(\tau) = \frac{\Delta(\tau)^2}{\Delta(2\tau)\Delta(\tau/2)} - 24, \tag{3.34}$$

where $\Delta = q\prod_{n=1}^{24}(1-q^n)^{24}$ is the discriminant, a modular form of weight 24. In (3.34), we have subtracted the constant 24 so that the expansion of $K(\tau)$ in powers of $q^{1/2}$ has no constant term:

$$\begin{aligned} K(\tau) &= q^{-1/2} + 276q^{1/2} + 2048q + 11202q^{3/2} + 49152q^2 + 184024q^{5/2} + 614400q^3 \\ &\qquad + 1881471q^{7/2} + 5373952q^4 + 14478180q^{9/2} + \dots . \end{aligned} \tag{3.35}$$

This is analogous to the definition of the J-function without a constant term. Another formula for K is

$$K = \frac{q^{-1/2}}{2}\left(\prod_{n=1}^{\infty}(1+q^{n-1/2})^{24} + \prod_{n=1}^{\infty}(1-q^{n-1/2})^{24}\right) + 2048q\prod_{n=1}^{\infty}(1+q^n)^{24}. \tag{3.36}$$

The product formula in (3.34), since it converges for all $|q| < 1$, shows that K is nonsingular as a function on $\mathcal{H}$. As for the behavior at the cusps, either formula shows that K has a simple pole at the NS cusp, that is, it behaves for $q \to 0$ as $q^{-1/2}$. K is regular at the Ramond cusp; in fact

$$K(\tau = 1) = -24. \tag{3.37}$$

This statement is equivalent to the statement that $K + 24 = \Delta(\tau)^2/\Delta(2\tau)\Delta(\tau/2)$ vanishes at $\tau = 1$. Indeed, that function has a pole at the NS cusp, so it must have a zero somewhere. Its representation as a convergent infinite product shows that it is nonzero for $0 < |q| < 1$, so the zero is at the Ramond cusp. We give another explanation of (3.37) in section 3.3 in analyzing the $k^* = 1$ model. The fact that the holomorphic function K on $\mathcal{Y}$ has only one pole, which is of first order, gives another way to prove that $\mathcal{Y}$ is of genus zero.

Now let us consider the Neveu-Schwarz partition function F of a holomorphic SCFT. Any Γ_θ-invariant function F on $\mathcal{H}$ can be written as a function of K. The function F arising in a holomorphic SCFT is actually *polynomial* in K. Indeed, since the definition of F as $\mathrm{Tr}\, q^{L_0}$ is convergent for $0 < |q| < 1$, F is regular as a function on $\mathcal{H}$. So the only poles of F are at cusps, but the formula (3.31), which reflects the fact that the ground state energy in the Ramond sector is zero, says that F has no pole at the Ramond cusp.

So the only pole of F is at the Neveu-Schwarz cusp, that is at $K = \infty$. Consequently, in any holomorphic SCFT, the Neveu-Schwarz partition function F is a polynomial in K.

The degree of this polynomial is precisely k^*, since $F \sim q^{-k^*/2}$ for $q \to 0$. So

$$F = \sum_{r=0}^{k^*} f_r K^r. \tag{3.38}$$

Thus, F depends on $k^* + 1$ coefficients.

Now, for the case of an SCFT that is dual to three-dimensional supergravity, what would we expect F to be? To follow the same logic as in the bosonic case, we would argue that since the minimum energy of a BTZ black hole is $L_0 = 0$, and the entropy vanishes at $L_0 = 0$, primary states corresponding to black holes should be absent for $L_0 \leq 0$. Equivalently, primary fields other than the identity should be absent for dimension less than $(k^* + 1)/2$. Either statement would mean that up to terms of order $q^{1/2}$, F would coincide with the naive function

$$F_0(k^*) = q^{-k^*/2} \prod_{n=2}^{\infty} \frac{1 + q^{n-1/2}}{1 - q^n} \tag{3.39}$$

that counts superconformal descendants of the identity.

For each positive integer k^*, there is a uniquely determined function F_{k^*} that is a polynomial in K and coincides with $F_0(k^*)$ up to order $q^{1/2}$. This is a natural analog of the partition function Z_k that we defined for an extremal CFT without supersymmetry.

However, the same logic that makes us think that there should be no NS primaries of $L_0 \leq 0$ also would make us believe that there should be no Ramond primaries of $L_0 = 0$. (There are automatically no Ramond states of $L_0 < 0$, since $L_0 = \mathcal{G}_0^2$.) In either the NS or Ramond sector, the classical black hole entropy vanishes for $L_0 \leq 0$.

The number of Ramond primaries of $L_0 = 0$ in a theory with Neveu-Schwarz partition function F_{k^*} is, according to (3.31), $h_0 = (-1)^{k^*} F_{k^*}(1)$. Let us call this number β_{k^*}. β_{k^*} is uniquely determined for each k^*. A practical way to determine it, using (3.37), is to write

$$\beta_{k^*} = (-1)^{k^*} F_{k^*}(1) = (-1)^{k^*} \sum_{r=0}^{k^*} f_r K(1)^r = (-1)^{k^*} \sum_{r=0}^{k^*} f_r (-24)^r. \tag{3.40}$$

So one first evaluates the coefficients f_r in (3.38) to ensure that there are no NS primary states with $L_0 \leq 0$, and then one evaluates the sum in (3.40).

The first ten values of β_{k^*} are given in Table 1. The first observation is that β_{k^*} never vanishes in this range. Consequently, it is impossible to assume that there are no primary operators other than the identity of dimension less than $(k^*+1)/2$. If we assume that there are no such primaries in the NS sector, then there are primaries of dimension $k^*/2$ in the Ramond sector.

k^*	β_{k^*}
1	24
2	24
3	95
4	1
5	143
6	1
7	262
8	-213
9	453
10	-261

Table 1. Values of the index β_{k^*} for $k^* = 1, \ldots, 10$.

The good news is that the numbers in the table are rather small compared to black hole multiplicities that arise in the classically allowed region $L_0 > 0$. So an optimistic view is to interpret the numbers in the table as quantum corrections. For example, at $k^* = 9$, where β takes the relatively large value 453, the multiplicity of the lowest mass classically allowed black hole, namely $L_0 = 1/2$ in the NS sector, turns out to be 135149371 if the partition function F_9 can be trusted. (As usual, we count only primaries, although this involves only a small correction.) The lowest classically allowed black hole in the Ramond sector, at $L_0 = 1$, has multiplicity 381161020987.

However, it may not seem logical to assume that there are no NS primaries at dimension $k^*/2$ and allow Ramond primaries of that dimension. Moreover, it really does not make sense to do this, since some values of β in the table are negative.

So we retreat from claiming that there are no NS primary fields of dimension less than $(k^*+1)/2$, and instead we consider the hypothesis that there are no such primaries of dimension less than $k^*/2$. In this case, we are free to add an integer s to F_{k^*}. If we do so, the number of NS primaries of dimension $k^*/2$ becomes $f_0 = s$, and the number h_0 of Ramond primaries of that dimension changes by $(-1)^{k^*}s$. The combination

$$\beta_{k^*} = h_0 - (-1)^{k^*} f_0 \tag{3.41}$$

is unchanged, and so is still given by the numbers in the table.

Unfortunately, then, we have no way to determine the numbers f_0 and h_0 separately, only the "index" β. We also have little insight about the more conventional index χ defined in (3.23). Both of these quantities characterize the primary states of $L_0 = 0$.

As in the bosonic case, one can construct useful formulas for the functions F_{k^*} in terms of Hecke operators acting on F_1. However, we will omit this.

Optimistic Conjecture

We define an extremal SCFT to be one with no primary fields other than the identity of dimension less than $k^*/2$. Equivalently, the NS partition function is equal up to an additive integer to the function F_{k^*} that was defined above.

Generalizing what we have said about the bosonic case, the most optimistic conjecture we can propose is that an extremal SCFT exists and is unique for every positive integer k^*, and is dual to three-dimensional supergravity. The best evidence we can offer is that extremal SCFT's in this sense do exist for $k^* = 1, 2$ and there are results about uniqueness [31,35] at least for $k^* = 1$.

In addition, as we will see, the functions F_{k^*} with $k^* = 3, 4$ have interesting properties suggestive of the existence of actual theories.

These matters are discussed in section 3.3. In addition, in the appendix, we describe the functions F_{k^*} for $5 \leq k^* \leq 10$.

3.3. Extremal SCFT's With Small k^*

k^*=1

The first construction of an extremal SCFT was made by FLM at $k^* = 1$, that is $c = 12$. They considered eight free bosons X_i compactified using the E_8 root lattice, combined with eight free fermions ψ_i to achieve superconformal symmetry. This theory has NS primary fields of dimension 1/2, namely the ψ_i. To eliminate these fields, they considered a $\mathbb{Z}_2$ orbifold, dividing by the operation that acts as -1 on all X_i and ψ_i.

They conjectured that this construction gave the unique SCFT that has $c = 12$ and no NS primary of dimension 1/2.

While the construction is simple, it has the drawback of not making manifest the global symmetry of the model. This was remedied much more recently by Duncan [35], who described the same model in another way. In this construction, one begins with 24 free fermions λ_i, $i = 1, \ldots, 24$, forming again a system with $c = 12$. In quantization, one can require the λ_i to be either antiperiodic or periodic around the spatial direction, giving what we will call NS_0 and R_0 sectors. (We reserve the name NS and R for another construction that will appear shortly.)

With 24 fermions, the ground state energy in the NS_0 sector is $-1/2$, while in the R_0 sector, the ground state energy is $+1$. The minimum dimension of a spin field is the difference between these two energies, or 3/2.

The model has an $O(24)$ symmetry that rotates the fields λ_i. The ground state in the R_0 sector is highly degenerate, because of zero modes of the fields λ_i. It consists of $2^{12} = 4096$ states, transforming in the spinor representation of $O(24)$ (or rather its double cover).

The spin fields of lowest dimension are therefore 4096 fields W_α of dimension 3/2 transforming as spinors. Dimension 3/2 is the correct dimension for a supercurrent, but a generic linear combination $W = \sum_\alpha \epsilon^\alpha W_\alpha$, with c-number coefficients ϵ^α, does not generate a superconformal algebra. Schematically, and without worrying about the precise coefficients, the operator product $W \cdot W$ gives

$$W(x)W(0) \sim \frac{\overline{\epsilon}\epsilon}{x^3} + \frac{\overline{\epsilon}\epsilon\, T}{x} + \frac{\overline{\epsilon}\Gamma^{ij}\epsilon\, \lambda_i\lambda_j}{x^2} + \frac{\overline{\epsilon}\Gamma^{ijkl}\epsilon\, \lambda_i\lambda_j\lambda_k\lambda_l}{x} + \text{regular}. \tag{3.42}$$

(Here T is the stress tensor and Γ_i are the gamma matrices of $O(24)$.) The condition for W to generate a superconformal algebra is that the last two terms should be absent. This is equivalent to

$$\overline{\epsilon}\Gamma^{ij}\epsilon = \overline{\epsilon}\Gamma^{ijkl}\epsilon = 0. \tag{3.43}$$

It is shown in [35] that a spinor ϵ obeying these conditions exists and, if normalized to have $\overline{\epsilon}\epsilon = 1$, is unique up to an $O(24)$ transformation. Another important fact is that any such ϵ has one definite $SO(24)$ chirality or the other.

Any choice of a solution of (3.43) turns the theory into an $\mathcal{N} = 1$ superconformal field theory, with W as the supercurrent. However, as ϵ is not $O(24)$-invariant, the theory regarded as an $\mathcal{N} = 1$ SCFT does not have $O(24)$ symmetry. Rather, the choice of ϵ breaks

$O(24)$ symmetry to a group that is known as the Conway group Co_0. It is the symmetry group of the Leech lattice, and is a double cover of a sporadic finite group Co_1. Thus, as a superconformal field theory, this model has symmetry Co_0, not $O(24)$.

With this choice of superconformal algebra, we need to understand what are the Neveu-Schwarz and Ramond vertex operators. A Ramond vertex operator $\mathcal{O}$ has a square root singularity in the presence of the supercurrent W:

$$W(x)\mathcal{O}(x') \sim \frac{\mathcal{O}'}{(x-x')^{n-1/2}} \tag{3.44}$$

for some integer n and some operator $\mathcal{O}'$. With W understood as a spin operator with respect to the original fermions λ, those fermions have precisely this property. So they are Ramond fields. This enables us to analyze all of the states in the original NS_0 sector. The ground state corresponds to the identity operator, whose OPE with W of course has no branch cut, so it is an NS rather than Ramond operator. The operators in the NS_0 sector that have no branch cut with W are those that are products of an even number of λ's and their derivatives. The partition function that counts the corresponding states is

$$\frac{q^{-1/2}}{2}\left(\prod_{n=1}^{\infty}(1+q^{n-1/2})^{24}+\prod_{n=1}^{\infty}(1-q^{n-1/2})^{24}\right). \tag{3.45}$$

We recognize this as part of the formula (3.36) for the function K. In the R_0 sector, of the 4096 ground states, half have one chirality or fermion number and half have the other. Any excitation at all can be combined with a ground state of properly chosen chirality to get an operator that either does or does not have a branch cut with W, as desired. So for each L_0 eigenvalue in the R_0 sector, precisely half the states contribute to the NS sector and half to the R sector. The contribution of R_0 states to the NS sector is therefore

$$2048q\prod_{n=1}^{\infty}(1+q^n)^{24}. \tag{3.46}$$

Adding up (3.45) and (3.46), we see that the total partition function F_1 of the NS sector in this model is precisely what we have called K:

$$F_1 = K = \frac{q^{-1/2}}{2}\left(\prod_{n=1}^{\infty}(1+q^{n-1/2})^{24}+\prod_{n=1}^{\infty}(1-q^{n-1/2})^{24}\right)+2048q\prod_{n=1}^{\infty}(1+q^n)^{24}. \tag{3.47}$$

We can similarly compute the Ramond partition function H_1 of this model. The contribution of the NS_0 sector is obtained from (3.45) by changing a sign so as to project onto states of odd fermion number, rather than even fermion number. And the contribution of the R_0 sector is the same as (3.46). So

$$\begin{aligned} H_1 &= \frac{q^{-1/2}}{2}\left(\prod_{n=1}^{\infty}(1+q^{n-1/2})^{24}-\prod_{n=1}^{\infty}(1-q^{n-1/2})^{24}\right)+2048q\prod_{n=1}^{\infty}(1+q^n)^{24} \\ &= 24+4096q+98304q^2+1228800q^3+10747904q^4+\dots. \end{aligned} \tag{3.48}$$

Except for states of $L_0=0$, the global supercharge $\mathcal{G}_0$ of the Ramond sector exchanges the part of the Ramond sector coming from NS_0 with the part coming from R_0. This implies that we can alternatively write

$$H_1 = 24+4096q\prod_{n=1}^{\infty}(1+q^n)^{24}, \tag{3.49}$$

where we have removed the NS_0 contribution except for the ground states, and doubled the R_0 contribution. The equivalence of these formulas is not very obvious.

A consequence of the above formulas is that h_0, the number of Ramond states of $L_0=0$, is equal to 24. This gives another explanation for (3.37). Another consequence of (3.49) and (3.47) is that for $L_0>0$, the coefficients h_n in the q-expansion of H_1 are precisely twice the corresponding coefficients f_n in the expansion of F_1:

$$h_n = 2f_n. \tag{3.50}$$

(Here h_n is defined for integer n, while f_n is defined for integer or half-integer n, so this formula involves all coefficients h_n but only half of the f_n.) This is highly exceptional for SCFT's and reflects the particular way that this one was constructed. The relation (3.50) is equivalent to

$$0 = F_1(\tau)+F_1(\tau+1)-H_1(\tau) = F_1(\tau)-G_1(\tau+1)-H_1(\tau). \tag{3.51}$$

Although disguised in our way of presenting it, this is a standard relation in string theory. In the FLM construction of the model via 8 free bosons and 8 free fermions, (3.51) amounts to the statement of Gliozzi, Olive, and Scherk [52] that the Ramond-Neveu-Schwarz model has equal partition function in the Ramond and Neveu-Schwarz sectors, because of spacetime supersymmetry.

The factor of 2 in (3.50) is consistent with asymptotic equality between the numbers of NS and Ramond black holes because it only involves half of the coefficients of F_1. H_1 has only half as many coefficients as F_1, but they are twice as big.

As we have explained, the NS partition function F_{k^*} of an extremal SCFT with any k^* is a polynomial in K or equivalently in F_1,

$$F_{k^*} = f(F_1), \tag{3.52}$$

for some polynomial f. From (3.29), it follows that the Ramond partition function H_{k^*} of an extremal SCFT can be obtained from H_1 using the same polynomial:

$$H_{k^*} = (-1)^{k^*} f(-H_1). \tag{3.53}$$

A Model With $k^* = 2$

What in our language is an extremal SCFT of $k^* = 2$ was constructed [26] by Dixon, Ginsparg, and Harvey (DGH), relatively soon after the work of FLM, by twisting the FLM monster construction in a somewhat similar fashion. Note that an extremal SCFT of $k^* = 2$ has $c = 24$, like the FLM monster theory.

We recall that the starting point of the FLM construction is 24 free bosons X_i that are compactified via the Leech lattice. Then one performs a $\mathbb{Z}_2$ orbifold, dividing by the symmetry $X_i \to -X_i$. To construct an orbifold, one first constructs untwisted and twisted sectors (by quantizing fields $X_i(\sigma)$ that are assumed to be periodic or antiperiodic functions of σ) and then one projects both sectors onto their $\mathbb{Z}_2$-invariant subspaces.

DGH observed that the ground state energy in the untwisted sector is -1, while the ground state energy in the twisted sector is $+1/2$. So a twist operator of lowest energy has dimension $3/2$, the right dimension for a supercurrent. They went on to show that it is possible to pick a twist operator $\mathcal{S}$ that generates a superconformal algebra.

In the FLM construction, the field $\mathcal{S}$ is not present, since it is odd rather than even under $\mathbb{Z}_2$, and is projected out when one forms the orbifold theory. However, DGH showed that it is possible to obtain an $\mathcal{N} = 1$ SCFT by modifying the usual orbifold projection. They defined the NS sector to consist of fields that do not have a cut in their OPE with $\mathcal{S}$, while the Ramond sector consists of fields that do have such a cut. Concretely, the NS sector consists of the $\mathbb{Z}_2$-even part of the untwisted sector of the orbifold plus the part of the twisted sector that transforms under $\mathbb{Z}_2$ as does $\mathcal{S}$ (so that $\mathcal{S}$, in particular, is an NS field), while the Ramond sector consists of the $\mathbb{Z}_2$-odd part of the untwisted sector plus the $\mathbb{Z}_2$-even part of the twisted sector.

From this information, one can of course compute the NS and Ramond partition functions F_2 and H_2. From our point of view, of course, F_2 is a polynomial in $F_1 = K$, chosen so that $F_2 = q^{-1} + \mathcal{O}(q^{1/2})$. This gives $F_2 = K^2 - 552$, and similarly, in view of (3.53), $H_2 = H_1^2 - 552$. From this, we get

$$\begin{aligned} F_2 &= q^{-1} + 4096q^{1/2} + 98580q + 1228800q^{3/2} + 10745856q^2 + 74244096q^{5/2} + 432155586q^3 + \dots \\ H_2 &= 24 + 196608q + 21495808q^2 + 864288768q^3 + \dots . \end{aligned} \tag{3.54}$$

The number of states of $L_0 = 0$ is 24 in the Ramond sector and 0 in the NS sector.

The global symmetry of the DGH model is closely related to the symmetry group Co_0 of the Leech lattice (it is an extension of this by a finite abelian group that involves the momenta of the Leech lattice). Since the $k^* = 1$ model also has Co_0 symmetry, one might optimistically conjecture that this group is relevant to supergravity at all k^*. However, results below about $k^* = 3$ indicate that this is not the case.

Partition Function With $k^* = 3$

Unfortunately, this is the last case for which a suitable SCFT is known, but we can of course determine the appropriate NS and Ramond partition functions for all k^*, at least modulo an additive integer. We consider $k^* = 3, 4$ here because they turn out to have unusual properties, and relegate the further cases $k^* = 5, 6, \dots, 10$ to an appendix.

For $k^* = 3$, we have

$$F_3 = K^3 - 828K - 6143. \tag{3.55}$$

This leads to

$$\begin{aligned} F_3 &= q^{-3/2} + 1 + 33606q^{1/2} + 1843200q + 43434816q^{3/2} + 648216576q^2 \\ &\quad + 7171304841q^{5/2} + 63903727616q^3 + \dots \\ H_3 &= 95 + 3686400q + 1296433152q^2 + 127807455232q^3 + \dots . \end{aligned} \tag{3.56}$$

Here one finds an identity just like (3.50): the expansion coefficients h_n of H_3 are related for $n > 0$ to analogous coefficients f_n of F_3 by

$$h_n = 2f_n. \tag{3.57}$$

This expresses the fact that

$$F_1^3(\tau) + F_1^3(\tau + 1) = H_1^3(\tau) - 1536, \tag{3.58}$$

along with the similar linear relation (3.51).

It seems that the identity $h_n = 2f_n$, $n > 0$, does not hold for any values of k^* except 1 and 3. In general, the differences $h_n - 2f_n$ are relatively small, reflecting the fact that asymptotically an NS or Ramond black hole has the same entropy, but they are not zero. The fact that these degeneracies are actually equal for $k^* = 3$ may be a clue to the construction of this model.

One important point about the $k^* = 3$ model is that it seems very unlikely to have Co_0 or Co_1 symmetry. Indeed the leading coefficient 95 of H_3 is not in any economical way the dimension of a representation of Co_0. (The dimensions of the first few irreducible representations of this group are 1, 24, 276, 299, and 1771; for Co_1, one must omit the number 24 from this list.) Adding an integer n to F_3 and therefore subtracting n from H_3, does not help; we cannot pick n so that n and $95 - n$ are both dimensions of Co_0 representations in an economical fashion. So it appears that, despite the tempting evidence from the cases $k^* = 1, 2$, the group Co_0 is probably not a general symmetry of three-dimensional supergravity.

Monster Symmetry For $k^* = 4$?

However, in the next case, $k^* = 4$, there may well be a much larger discrete symmetry group, namely the monster group $\mathbb{M}$.

For $k^* = 4$, we have

$$F_4 = K^4 - 1104K^2 - 8191K + 107545. \tag{3.59}$$

This leads to the NS and Ramond partition functions

$$\begin{aligned} F_4 &= q^{-2} + q^{-1/2} + 1 + 196884q^{1/2} + 21493760q + 864299970q^{3/2} + 20246053140q^2 \\ &\quad + 333202640600q^{5/2} + 4252023300096q^3 + 44656994071935q^{7/2} + 401490908149760q^4 + \dots \\ H_4 &= 1 + 42987520q + 40491712512q^2 + 8504046600192q^3 + 802981773312000q^4 + \dots . \end{aligned} \tag{3.60}$$

The first non-trivial coefficient in F_4 is the famous number $196884 = 196883 + 1$ whose appearance in the J function gave the original hint of a connection between $\mathbb{M}$ and modular functions and ultimately conformal field theory. Moreover, the leading coefficient in H_4 is 1, which is compatible with any assumption about an automorphism group, since it is the dimension of an irreducible representation, namely the trivial one. The contrasts with what we found for $k^* = 3$, where the leading coefficient, namely 95, makes it difficult to postulate a large discrete symmetry.

To explore further the hypothesis of $\mathbb{M}$ symmetry, we attempt to express the coefficients in F_4 and H_4 in terms of dimensions of representations of $\mathbb{M}$. The dimensions $d_1, d_2, \ldots, d_{12}$ of the first 12 monster representations, which we call R_i, $i = 1, \ldots, 12$, are given in the table.

d_1	1
d_2	196883
d_3	21296876
d_4	842609326
d_5	18538750076
d_6	19360062527
d_7	293553734298
d_8	3879214937598
d_9	36173193327999
d_{10}	125510727015275
d_{11}	190292345709543
d_{12}	222879856734249

Table 2. Presented here from [53] are the dimensions d_i of the i^{th} irreducible representation of the monster group $\mathbb{M}$, for $i = 1, \ldots, 12$. We denote as R_i the representation of dimension d_i.

A little experimentation soon shows that the nontrivial coefficients in the NS partition function F_4 can be nicely written in terms of the d_i. The first 8 nontrivial coefficients $f_{1/2}, f_1, \ldots, f_4$ are

$$\begin{aligned}
196884 &= d_1 + d_2 \\
21493760 &= d_1 + d_2 + d_3 \\
864299970 &= 2d_1 + 2d_2 + d_3 + d_4 \\
20246053140 &= 3d_1 + 4d_2 + 2d_3 + d_4 + d_6 \\
333202640600 &= 4d_1 + 5d_2 + 3d_3 + 2d_4 + d_5 + d_6 + d_7 \\
4252023300096 &= 5d_1 + 7d_2 + 4d_3 + 4d_4 + 2d_5 + 2d_6 + d_7 + d_8 \\
44656994071935 &= 7d_1 + 11d_2 + 7d_3 + 6d_4 + 3d_5 + 4d_6 + 2d_7 + 2d_8 + d_9 \\
401490908149760 &= 10d_1 + 16d_2 + 12d_3 + 9d_4 + 5d_5 + 7d_6 + 4d_7 + 4d_8 + d_9 + d_{10} + d_{12}.
\end{aligned} \tag{3.61}$$

The coefficients on the right hand side are rather small given the numbers involved, and as we will explain shortly, they do not grow much faster than is required by superconformal symmetry. The first three numbers here equal the first three non-trivial coefficients of the J function, but afterwards the two series diverge.

The coefficients h_n in the q-expansion of H_4 can similarly be expressed in terms of dimensions of monster representations. In doing so, however, the following is a useful shortcut and also possibly a clue to constructing the model. First of all, in contrast to the case of $k^* = 1, 3$, it is not true that $h_n - 2f_n = 0$ for $k^* = 4$ and all $n > 0$. But some of these differences vanish, and they are all relatively small. As we have already expressed the f_n for small n in terms of monster dimensions in (3.61), it suffices now to find similar expressions for $h_n - 2f_n$. Actually, as the h_n are all even for $n > 0$, we prefer to divide by 2, and we find

$$\begin{aligned}
\frac{h_1}{2} - f_1 &= 0 \\
\frac{h_2}{2} - f_2 &= -f_{1/2} \\
\frac{h_3}{2} - f_3 &= 0 \\
\frac{h_4}{2} - f_4 &= -f_1 \\
\frac{h_5}{2} - f_5 &= 0.
\end{aligned} \tag{3.62}$$

When this is combined with (3.61), we see that the $h_i/2$, $i \leq 5$ are linear combinations of the d_i with small positive integer coefficients. So the initial coefficients of the Ramond partition function are compatible with $\mathbb{M}$ symmetry, and in addition many of the differences $h_n/2 - f_n$ vanish.

T. Gannon has given a simple explanation of these results by showing that one can express F_4 and H_4 in terms of the J-function via $F_4(\tau) = J(2\tau) + J(\tau/2) + 1$, $H_4(\tau) = J(\tau/2) + J((\tau+1)/2) + 1$. Hence F_4 and H_4 inherit a relation to the monster from J. The fact that the formulas have such a direct explanation may lessen the case for a new SCFT with monster symmetry.

Given that the above formulas exist, they are not quite unique, because there are some linear relations[17] among the d_i with small coefficients:

$$\begin{aligned} d_1 + d_4 + d_5 &= d_3 + d_6 \\ d_3 + d_8 + d_{12} &= d_2 + d_7 + d_9 + d_{11}. \end{aligned} \tag{3.63}$$

We have made some choices to ensure that the above formulas are compatible with superconformal symmetry, in the following sense. The first formula in (3.61) indicates that at $L_0 = 1/2$, the NS sector has 196883 primary states $|\rho_i\rangle$ transforming in the representation R_2. It follows that at $L_0 = 1$, there are descendants $\mathcal{G}_{-1/2}|\rho_i\rangle$ transforming in the same representation. This continues at higher levels; the number of copies of R_2 appearing as descendants of the primary states $|\rho_i\rangle$ at $L_0 = s/2$ is $1, 1, 1, 2, 3, 4, 5, 7$ for $s = 1, 2, \ldots, 8$. If we count only primary states, we find, assuming that (3.61) is the right decomposition, that primary states in the representation R_2 occur at $L_0 = 1/2, 3/2, 2, 7/2, 4$, each time with multiplicity 1. Similar remarks apply to other representations. A special case is that, for the range of L_0 considered in (3.61), most of the states that transform in the trivial representation R_1 are actually descendants of the identity. If (3.61) is the right decomposition, then the first primary field that is $\mathbb{M}$-invariant is at $L_0 = 5/2$. If we were to rewrite (3.61) in terms of primary states only, the coefficients would be even smaller (but still nonnegative), strengthening the case for monster symmetry.

In this particular example, we cannot add an integer to F_4 without spoiling the hypothesis of monster symmetry. If there were an NS primary with $L_0 = 2$, it would have a descendant at $L_0 = 5/2$, and the number of primaries at $L_0 = 5/2$ would be less than 196883.

In the appendix, we present somewhat similar though less extensive evidence for baby monster symmetry at $k^* = 6$.

Sum Over Spin Structures

---

[17] These relations can be explained as follows. The symmetric part of $R_2 \otimes R_2$ decomposes as $\mathrm{Sym}^2 R_2 \cong R_1 \oplus R_2 \oplus R_4 \oplus R_5$, and the antisymmetric part decomposes as $\wedge^2 R_2 \cong R_3 \oplus R_6$. On the other hand, $\dim(\mathrm{Sym}^2 R_2) - \dim(\wedge^2 R_2) = \dim R_2 = d_2$. Taken together, these facts imply the first relation in (3.63), and the second follows similarly by considering the decomposition of $R_3 \otimes R_3$.

Given an SCFT with c an integer multiple[18] of 24, by summing over spin structures, one can make an ordinary bosonic CFT. The sum over spin structures projects both the NS and Ramond sectors onto bosonic states of integer dimension, as is appropriate for a bosonic CFT.

For c to be a multiple of 24, k^* must be even. It was already shown by DGH that the sum over spin structures, applied to their model which in our language is $k^* = 2$, gives the FLM model at $k = 1$. A natural question (raised by J. Duncan) is what this operation does at higher even values of k^*.

The sum over spin structures removes fermionic operators such as the supercurrent $\mathcal{S}$, of dimension 3/2. However, it leaves operators such as $\mathcal{S}\partial\mathcal{S}$ that are even in $\mathcal{S}$. This operator has dimension 4, and is a Virasoro primary (though, of course, it is a descendant in the $\mathcal{N} = 1$ super-Virasoro algebra). A bosonic extremal CFT of $k \geq 4$ should not have a primary of dimension 4. Hence, the "sum over spin structures" operation applied to an extremal SCFT of $k^* \geq 8$ does not give an extremal CFT. However, it does give an extremal CFT if applied to a theory of $k^* = 4$ or 6.

Hence, if extremal SCFT's exist with $k^* = 4, 6$, they can be used to generate extremal CFT's of $k = 2, 3$. It is interesting that $k^* = 4, 6$ are precisely the values at which extremal SCFT's may have monster or baby monster symmetry. (The baby monster group is the centralizer of an involution in the monster, and hence it is conceivable for an SCFT at $k^* = 6$ to have baby monster symmetry while the CFT at $k = 3$ has monster symmetry. The framework for this is described in [26].)

Part of what we have said can be understood in another way. Supergravity and gravity are different in the semiclassical limit, and one cannot be obtained from the other by merely summing over spin structures. So whatever is the SCFT dual of supergravity (even if our assumptions are too optimistic and it is not extremal), the sum over spin structures applied to this SCFT cannot give, for arbitrarily large k^*, the CFT that is dual to gravity.

---

[18] If c is an odd multiple of 12, the sum over spin structures is modular-invariant but projects out the NS ground state and does not give a CFT in the usual sense. This is actually important in superstring theory, where, with $c = 12$ in the light-cone treatment, the NS ground state is a "tachyon" and is removed in the sum over spin structures.

4. $k = 2$ Partition Function On A Hyperelliptic Riemann Surface

The most important question raised by this paper is certainly whether appropriate CFT's and SCFT's exist, beyond the few examples that are known.

In this section, we will perform a small computation that aims to give a hint that this is the case, at least for the bosonic theory with $k = 2$ and hence $c = 48$. We will show that the partition function of such a theory on a hyperelliptic Riemann surface of any genus can be determined in a unique and consistent way. This includes, for example, any Riemann surface of genus 2. The fact that we get a unique and consistent result in this situation gives some encouragement for believing that the $k = 2$ model does exist and is unique.

In this paper, we merely demonstrate the consistency of an algorithm for determining the partition function. Hopefully it will be possible in future work to get explicit formulas, at least for genus 2. Our method also works for $k = 1$, though here the ability in principle to determine the partition function comes as no surprise, since the model has been constructed explicitly [24] and in fact the genus 2 partition function has been computed [54] by a quite different method.

Perhaps it would help orient the reader to compare what we will do to another possible approach to determining the partition function. Genus 2 partition functions were determined in [54] for a variety of holomorphic CFT's with $c = 24$. The basic method was to express the partition function in terms of a Siegel modular form, which depends on only finitely many coefficients, and determine the coefficients by considering the behavior when the Riemann surface C degenerates. There are two types of degeneration (C can break up into two genus 1 curves joined at a point, or can reduce to a single genus 1 curve with two points glued together). The partition function can be determined from the behavior at just one degeneration, and there is a problem of consistency to show that one gets the same result either way. One could attempt to demonstrate the consistency by studying the appropriate Siegel modular forms, but we prefer to prove consistency by establishing the associativity of a certain operator product algebra. One advantage is that this method can be used to determine the partition function on a hyperelliptic Riemann surface of any genus. On the other hand, if one could overcome the technicalities in genus 2 (see [55] for one approach), one could possibly compute the genus 2 partition function of an extremal CFT for all k.

4.1. Twist Fields

A hyperelliptic Riemann surface C is a double cover of the complex plane, for example a double cover of the complex x-plane, which we will call C_0, described by an equation

$$y^2 = \prod_{i=1}^{2g+2} (x - e_i). \tag{4.1}$$

The $2:1$ cover $C \to C_0$ is branched at the points $e_1, \ldots, e_{2g+2}$. C is smooth if the e_i are distinct, and has genus g if the number of branch points is precisely $2g+2$.

Our approach to determining the partition function on such a Riemann surface is based on an old idea [56-59]. The partition function of a conformal field theory $\mathcal{W}$ on the hyperelliptic Riemann surface C can be determined by computing, in a doubled theory, the correlation function of $2g+2$ copies of a "twist field" $\mathcal{E}$, inserted at the points $e_1, \ldots, e_{2g-2}$ in C_0.

Consider any CFT $\mathcal{W}$ of central charge c. Away from branch points, the theory $\mathcal{W}$ on the double cover C looks locally like the theory $\mathcal{W} \times \mathcal{W}$ on C_0. Here we have one copy of $\mathcal{W}$ for each of the two branches of $C \to C_0$.

In going around a branch point, the two copies of $\mathcal{W}$ are exchanged. So a more complete description is to say that the theory on C_0 is $\mathrm{Sym}^2\mathcal{W}$, the symmetric product of two copies of $\mathcal{W}$. The symmetric product theory is an orbifold of the product theory $\mathcal{W} \times \mathcal{W}$ in which one divides by the $\mathbb{Z}_2$ symmetry that exchanges the two branches.

Let us describe what kind of operators the symmetric product theory has. First, there are "untwisted" operators. These are simply operators of the theory $\mathcal{W} \times \mathcal{W}$ that are invariant under the exchange of the two factors. For example, let T^+ and T^- be the stress tensors of the two factors. Any local operator constructed as a polynomial in T^+ and T^- and their derivatives and invariant under the exchange $T^+ \leftrightarrow T^-$ gives an operator in the symmetric product theory. Operators of this particular type generate a chiral algebra that we will call $\mathrm{Sym}^2\mathcal{V}$, where $\mathcal{V}$ denotes the chiral algebra generated by a single stress tensor.

In addition, there are "twisted sector" operators. These operators correspond (in the operator/state correspondence of CFT) to states obtained by quantizing the theory $\mathcal{W} \times \mathcal{W}$ on a circle S^1, in such a way that the two copies of $\mathcal{W}$ are exchanged in monodromy around the circle. These states are the states of a *single* copy of $\mathcal{W}$ on a circle of twice the circumference. The energies of the states in the twisted sector of the $\mathrm{Sym}^2\mathcal{W}$ theory are therefore precisely one-half the energies of the original $\mathcal{W}$ theory (in conformal field theory, doubling the circumference of the circle divides the Hamiltonian by two).

Operators related to states in the twisted sector are called twist fields. The dimension of a twist field is the same as the difference in energy between the corresponding twisted sector state and the untwisted ground state. The ground state energy in the untwisted sector is $-2 \cdot c/24$ (where the 2 comes from the two copies of $\mathcal{W}$) and the ground state energy in the twisted sector is $-(1/2) \cdot c/24$ (where the factor of 1/2 was explained in the last paragraph). The difference is $d_{\mathcal{E}} = (c/24)(-1/2 - (-2)) = (3/2)c/24$, and this is the dimension of the twist field $\mathcal{E}$ of lowest energy. For our application, we are interested in the case that $c/24$ is an integer k, and hence $d_{\mathcal{E}} = 3k/2$.

Now we can explain how the partition function of the theory $\mathcal{W}$ on the hyperelliptic Riemann surface C defined in eqn. (4.1) can be interpreted as a genus zero correlation function in the theory $\mathrm{Sym}^2\mathcal{W}$. From the standpoint of the $\mathrm{Sym}^2\mathcal{W}$ theory on the x-plane, the role of the branch points is just to exchange the two copies of $\mathcal{W}$, via a twist field. On the double cover, there is no operator insertion at the branch points except the identity; the identity operator at a branch point corresponds in the downstairs description to the ground state in the twisted sector and hence to the twist operator $\mathcal{E}$ of lowest dimension. So the partition function on the double cover can be expressed in terms of the correlation function $\left\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\ldots\mathcal{E}(e_{2g+2})\right\rangle$ on the x-plane. (The precise statement involves the conformal anomaly; this is deferred to section 4.5.)

To compute this correlation function, we need to understand the chiral algebra that is obtained by extending the symmetric product $\mathrm{Sym}^2\mathcal{V}$ of two Virasoro algebras by the field $\mathcal{E}$. For this, we need to know what primary fields (of $\mathrm{Sym}^2\mathcal{V}$) appear in the operator product $\mathcal{E} \cdot \mathcal{E}$. To calculate the product $\mathcal{E} \cdot \mathcal{E}$, we need to look at the behavior when a pair of branch points approach each other. This situation is described locally by an equation

$$y^2 = (x - e)(x - e'), \tag{4.2}$$

and we are interested in the behavior for $e \to e'$. In that limit, the equation becomes $y^2 = (x - e)^2$, and the surface breaks up into two branches $C_\pm : y = \pm(x - e)$. Each branch is a copy of the complex plane. We can express the product $\mathcal{E}(e) \cdot \mathcal{E}(e')$ for $e \to e'$ as a sum of operators of the form $\mathcal{U}_+ \otimes \mathcal{U}_-$, where $\mathcal{U}_\pm$ is an operator in the original CFT $\mathcal{W}$ on the branch $C_\pm$. Moreover, the $\mathcal{U}_\pm$ are descendants of some Virasoro primary fields $\mathcal{O}_\pm$.

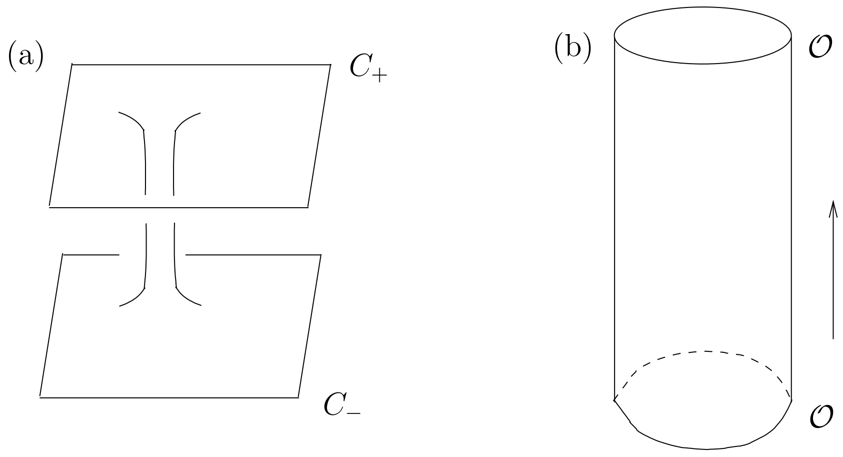


Fig. 5:

(a) The equation $y^2 = (x-e)(x-e')$ describes two branches connected by a tube. The tube collapses to a point as $e \to e'$. (b) Topologically, the tube is equivalent to a cylinder. Propagation along the tube multiplies any energy eigenstate by a c-number factor. If a primary field $\mathcal{O}$ flows in from the past, the same field flows out in the future. Alternatively, if we think of all fields as flowing out, then the two fields emerging at the two ends of the cylinder are conjugate.

The decomposition of $\mathcal{E}(e)\cdot\mathcal{E}(e')$ as a sum of descendants of operators $\mathcal{O}_+\otimes\mathcal{O}_-$ can be made in such a way that the operator $\mathcal{O}_+$ is conjugate to $\mathcal{O}_-$, meaning that the two-point function $\langle\mathcal{O}_+\mathcal{O}_-\rangle$ (in the original theory $\mathcal{W}$) is nonzero in genus zero. To see this, go back to the smooth double cover with $e \neq e'$. For e near e', the two branches are smoothly connected by a tube (fig. 5a). The surface thus has the topology of a cylinder (fig. 5b). Propagation along this cylinder is the "identity operator" on primary fields. So if a given primary state flows in at the bottom, it or flows out at the top. When we write $\mathcal{E}(e)\cdot\mathcal{E}(e')$ as a sum of descendants of operators $\mathcal{O}_+\otimes\mathcal{O}_-$, the convention is that both $\mathcal{O}_+$ and $\mathcal{O}_-$ are outgoing, so they are conjugate, rather than being equal. At any rate, the important conclusion for us is that $\mathcal{O}_+$ and $\mathcal{O}_-$ have the same dimension.

According to our interpretation of three-dimensional gravity, it is described by a conformal field theory with $c = 24k$ and no primary field other than the identity of dimension less than $k+1$. Hence, a primary operator $\mathcal{O}_+\otimes\mathcal{O}_-$ of the symmetric product theory, where $\mathcal{O}_+$ and $\mathcal{O}_-$ have the same dimension, is either the identity operator or has dimension at least $2(k+1)$.

To understand the chiral algebra or superalgebra[19] generated by $\mathcal{E}$ together with $\mathrm{Sym}^2\mathcal{V}$, we need to know all primary fields that appear with singular coefficients in the operator product $\mathcal{E}(e)\cdot\mathcal{E}(e')$. Since $\mathcal{E}$ has dimension $3k/2$, an operator that will contribute a singularity must have dimension less than $3k$ and hence at most $3k-1$. On the other hand, we have just seen that all operators appearing in the OPE are either descendants of the identity or have dimension at least $2(k+1)$.

For $k=1$ or 2, we have $3k-1<2(k+1)$. This leads to a drastic simplification, which is the reason that we will restrict ourselves here to $k=1,2$: singularities in the product $\mathcal{E}(e)\cdot\mathcal{E}(e')$ come only from the identity operator and its $\mathrm{Sym}^2\mathcal{V}$ descendants. Therefore, to understand correlation functions of $\mathcal{E}$, we need only understand the chiral algebra that is obtained by adding to $\mathrm{Sym}^2\mathcal{V}$ the primary field $\mathcal{E}$ of dimension $3k/2$ with the OPE schematically

$$\mathcal{E}\cdot\mathcal{E}\sim 1+\text{descendants}. \tag{4.3}$$

This is a chiral algebra that can be described in closed form and explicitly proved to obey the Jacobi identity. The only $\mathrm{Sym}^2\mathcal{V}$ primaries are 1 and $\mathcal{E}$. The only genus zero three point functions of primaries are $\langle 1\cdot 1\cdot 1\rangle$, which is trivial, and $\langle 1\cdot\mathcal{E}\cdot\mathcal{E}\rangle$, which is almost trivial in the sense that it reduces to a two-point function. To understand correlation functions of descendants, it suffices to understand the module for $\mathrm{Sym}^2\mathcal{V}$ consisting of $\mathcal{E}$ and its descendants. One way to construct this module explicitly is to observe that on the double cover of the x-plane, $\mathcal{E}$ simply corresponds to the identity operator and $\mathrm{Sym}^2\mathcal{V}$ to the ordinary Virasoro algebra; so this module for $\mathrm{Sym}^2\mathcal{V}$ can be deduced from the identity module for $\mathcal{V}$.

To get beyond the three-point function, the key step is to show that the chiral algebra obtained by adding $\mathcal{E}$ to $\mathrm{Sym}^2\mathcal{V}$ is consistent, that is, that the Jacobi identity is obeyed. If so, this chiral algebra determines all genus zero correlation functions of $\mathcal{E}$, and hence determines the partition function of our CFT on a general hyperelliptic Riemann surface.

In such a problem of trying to extend a known chiral algebra $\mathrm{Sym}^2\mathcal{V}$ by additional primary fields $\mathcal{E}$ (of integer or half-integer dimension, in which case one will get an extended chiral algebra or a superalgebra), the Jacobi identity is equivalent [60] to the statement that there exists a four-point function of the additional primary fields that has the appropriate singularities determined by the operator product expansion in all channels. (Such a function automatically has the right symmetries.) In the case at hand, since the only primary field that we are trying to add is $\mathcal{E}$, the only non-trivial primary four-point function that we have to consider is $\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$. In section 4.3, we explicitly construct this four-point function and show that it behaves correctly in all channels.

---

[19] As $\mathcal{E}$ has dimension $3k/2$, it is a fermionic operator if k is odd.

In the following sense, this result is not at all surprising. The four-point function $\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$ is essentially the partition function of the underlying CFT on the hyperelliptic Riemann surface $y^2 = (x-e_1)(x-e_2)(x-e_3)(x-e_4)$. That surface has genus 1, and we already know from section 3 that for every k there is a unique, natural genus 1 partition function with the properties we want. In section 4.4, we show explicitly, for $k = 1, 2$, that the four-point function $\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$ determined from the OPE's agrees with the genus 1 partition function as described in section 3.

4.2. The $\mathcal{E} \cdot \mathcal{E}$ Operator Product

As a first step in that direction, we will calculate the details of the $\mathcal{E} \cdot \mathcal{E}$ operator product. For this, we start with a double cover C of the x-plane branched at e and $e' = -e$, and so described by an equation $y^2 = x^2 - e^2$. If $u = x + y$, $v = x - y$, then the equation is $uv = e^2$. The two branches C_+ and C_- correspond respectively to $u \to \infty$, $v \to 0$ and $u \to 0$, $v \to \infty$.

The path integral over C gives a quantum state Ψ in the theory $\mathcal{W} \times \mathcal{W}$, that is, one copy of $\mathcal{W}$ for each branch. This state is invariant under exchange of the two branches by the symmetry $y \to -y$, so it is really a state in the symmetric product theory $\mathrm{Sym}^2\mathcal{W}$. We are really only interested in the part of Ψ proportional to the vacuum state and its descendants. As we have seen, this part suffices to describe the desired chiral algebra if $k \leq 2$.

We will determine Ψ by using the fact that certain elements in the product $\mathcal{V}\times\mathcal{V}$ of two Virasoro algebras annihilate Ψ. This is so because there are globally-defined holomorphic vector fields on C, of the form $V_n = 2^{-n}u^{n+1}d/du = -2^{-n}(e^2/v)^n v d/dv$. Let $\mathcal{S}$ be a contour on the surface C that wraps once around the "hole." If T is the stress tensor, the contour integral $\int_{\mathcal{S}} V_n T$ can be regarded, for any n, as an operator acting on the state Ψ. This operator is invariant under deformation of the contour. It can be deformed (fig. 6) to a contour $\mathcal{S}_+$ in the upper branch C_+ or a contour $\mathcal{S}_-$ in the lower branch C_-. So we have for all n

$$\left(\int_{\mathcal{S}_+} V_n T - \int_{\mathcal{S}_-} V_n T\right)\Psi = 0. \tag{4.4}$$

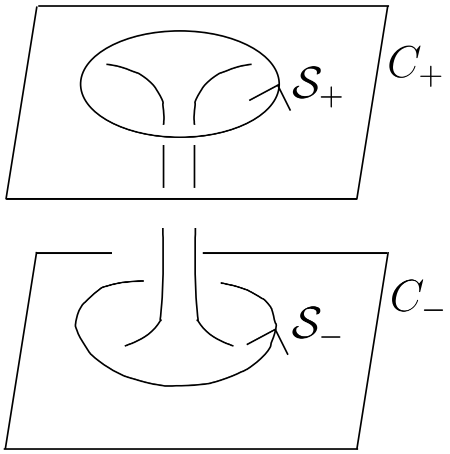


Fig. 6:

A contour $\mathcal{S}$ that wraps once around the hole can be deformed to the contour $\mathcal{S}_+$ in the upper branch or to the contour $\mathcal{S}_-$ in the lower branch.

We want to express the two contour integrals that appear here in terms of Virasoro generators on the two branches. To do this, we simply express V_n as a vector field on the branch C_+ or C_-, either of which we identify with the x-plane. For example, on the branch C_+, we write explicitly $y = \sqrt{1 - e^2/x^2} = 1 - e^2/2x^2 - e^4/8x^4 + \mathcal{O}(e^6)$. We have carried the expansion far enough to determine (for $k \leq 2$) all singular terms in the product $\mathcal{E}(e) \cdot \mathcal{E}(-e)$. So

$$V_n = x^{n+1}\left(1 - (n+2)\frac{e^2}{4x^2} + \frac{e^4}{x^4}\left(\frac{n^2+n-4}{32}\right) + \mathcal{O}(e^6)\right)\frac{d}{dx}. \tag{4.5}$$

This means, if we ignore the conformal anomaly for the moment, that $\int_{\mathcal{S}_+} V_n T$ corresponds, on the branch C_+, to the operator

$$Q_n^+ = L_n^+ - \frac{n+2}{4}e^2 L_{n-2}^+ + \left(\frac{n^2+n-4}{32}\right)e^4 L_{n-4}^+ + \mathcal{O}(e^6). \tag{4.6}$$

As a check, one can verify that $[Q_n^+, Q_m^+] = (n-m)Q_{n+m}^+$. Similarly, $\int_{\mathcal{S}_-} V_n T$ corresponds on the branch C_- to the operator

$$Q_n^- = \left(\frac{e^2}{4}\right)^n \left(L_{-n}^- - e^2\left(\frac{-n+2}{4}\right)L_{-n-2}^- + e^4\left(\frac{n^2-n-4}{32}\right)L_{-n-4}^- + \mathcal{O}(e^6)\right). \tag{4.7}$$

The state Ψ is determined for each value of e (up to multiplication by a complex scalar) by the condition that $\widehat{Q}_n\Psi = 0$, where $\widehat{Q}_n = Q_n^+ - Q_n^-$. However, because of the Virasoro anomaly some c-number terms must be added to the above formulas, reflecting the conformal anomaly in the mapping from u to x. There is no such c-number contribution to $\widehat{Q}_0$, since it cancels between the two branches. The c-numbers in $\widehat{Q}_n$ for other n can be conveniently determined by requiring that $[\widehat{Q}_n, \widehat{Q}_m] = (n-m)\widehat{Q}_{n+m}$. For our purposes, the only formulas we need are

$$\begin{aligned}
\widehat{Q}_0 &= \left(L_0^+ - \frac{e^2}{2}L_{-2}^+ - \frac{e^4}{8}L_{-4}^-\right) - \left(L_0^+ - \frac{e^2}{2}L_{-2}^- - \frac{e^4}{8}L_{-4}^-\right) + \dots \\
\widehat{Q}_1 &= \left(L_1^+ - \frac{3e^2}{4}L_{-1}^+ - \frac{e^4}{16}L_{-3}^+\right) - \frac{e^2}{4}\left(L_{-1}^- - \frac{e^2}{4}L_{-3}^-\right) + \dots \\
\widehat{Q}_2 &= L_2^+ - e^2L_0^+ - 3ke^2 + \frac{e^4}{16}L_{-2}^+ - \frac{e^4}{16}L_{-2}^- + \dots .
\end{aligned} \tag{4.8}$$

Terms of order e^6 have been omitted. The constant in $\widehat{Q}_2$ was obtained from $[\widehat{Q}_0, \widehat{Q}_2] = -2\widehat{Q}_2$.

By requiring that $\widehat{Q}_m\Psi = 0$ for $m = 0, 1, 2$, and that Ψ converges to the Fock vacuum $|\Omega\rangle$ for $e \to 0$, we now find Ψ to be

$$\begin{aligned}
\Psi(e) = \Bigg(1 &+ \frac{e^2}{4}(L_{-2}^+ + L_{-2}^-) + \frac{e^4}{32}(L_{-4}^+ + L_{-4}^-) \\
&+ \frac{e^4}{32}(L_{-2}^+ + L_{-2}^-)^2 + \frac{e^4}{192k}L_{-2}^+L_{-2}^- + \dots\Bigg)|\Omega\rangle.
\end{aligned} \tag{4.9}$$

We can immediately use this formula to determine the singular part of the $\mathcal{E}\cdot\mathcal{E}$ OPE. We normalize $\mathcal{E}$ so that the most singular term is $\mathcal{E}(x/2)\mathcal{E}(-x/2) \sim 1/x^{3k}$. Then, setting $e = x/2$, we get

$$\begin{aligned}
\mathcal{E}(x/2)\mathcal{E}(-x/2) \sim \frac{1}{x^{3k}}\Bigg(1 &+ \frac{x^2}{16}T + \frac{x^4}{2^{10}}\partial^2T \\
&+ \frac{x^4}{2^9}T\star T + \frac{x^4}{3k\cdot 2^{10}}T^+T^-\Bigg) + \mathcal{O}(x^{-3k+6}).
\end{aligned} \tag{4.10}$$

All we have done is to write, on the right hand side, the operator that corresponds to the state $\Psi(e)/x^{3k}$, for $e = x/2$. (The factor of $1/x^{3k}$ must be supplied by hand, since in defining $\Psi(e)$, we just normalized it so that the coefficient of the Fock vacuum is 1.) Also, T^+ and T^- are the stress tensors on the branches C_+ and C_-, respectively. $T = T^+ + T^-$ is the diagonal or total stress tensor, and similarly $L_n = L_n^+ + L_n^-$. In addition, we use the fact that in the correspondence between operators and states, the states $L_{-2}^{\pm}|\Omega\rangle$ correspond to the operators $T^{\pm}$, while $L_{-4}^{\pm}|\Omega\rangle$ correspond to $\frac{1}{2}\partial^2T^{\pm}$. Finally, we have written $T\star T$ for the operator corresponding to the state $L_{-2}^2|\Omega\rangle$. $T\star T(0)$ can be obtained as the term of order x^0 in the operator product $T(x)\cdot T(0)$.

4.3. Determining The Four-Point Function

Our goal is to use the explicit formula (4.10) to show that the chiral algebra obtained by adjoining $\mathcal{E}$ to $\mathrm{Sym}^2\mathcal{V}$ obeys the Jacobi identity. The Jacobi identity is equivalent to the existence of a four-point function for primary fields with the right symmetry properties and the right OPE singularities in all channels. It is enough to consider primary fields with respect to the $\mathrm{Sym}^2\mathcal{V}$ algebra, which we already know to exist. The only non-trivial case is the four-point function of $\mathcal{E}$.

For $k = 1$, the singular part of the $\mathcal{E} \cdot \mathcal{E}$ operator product only involves the diagonal stress tensor $T = T^+ + T^-$. This is a substantial simplification; the extended chiral superalgebra obtained by incorporating $\mathcal{E}$ can be viewed an extension of an ordinary Virasoro algebra $\mathcal{V}$, rather than a symmetric product $\mathrm{Sym}^2\mathcal{V}$. In fact, this algebra is a familiar one, an $\mathcal{N} = 1$ super Virasoro algebra:

$$\mathcal{E}(x/2)\mathcal{E}(-x/2) \sim \frac{1}{x^3}\left(1 + \frac{x^2}{16}T\right) + \mathcal{O}(x). \tag{4.11}$$

The stress tensor T has $c = 48$ (as it is the diagonal stress tensor in the symmetric product of two copies of a $c = 24$ theory). A superconformal field theory with this central charge can be constructed from 32 chiral multiplets consisting of bosonic and fermionic fields ϕ_i and ψ_i, where $\langle\phi_i(x)\phi_j(0)\rangle = -\delta_{ij}\ln(x)$, and $\langle\psi_i(x)\psi_j(0)\rangle = \delta_{ij}/x$, for $i, j = 1, \ldots, 32$. In terms of these fields, we can take

$$\mathcal{E} = \frac{i}{4\sqrt{2}}\sum_{k=1}^{32}\psi_k\partial\phi_k. \tag{4.12}$$

With this free field realization, all correlation functions of $\mathcal{E}$ can be explicitly computed. For example, the four-point function is

$$\begin{aligned}\big\langle\mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\big\rangle = \prod_{i=1}^{4}(de_i)^{3/2}\Bigg(&\left(\frac{1}{(e_1 - e_2)^3(e_3 - e_4)^3} + \text{cyclic}\right)\\ &+\frac{3}{32}\frac{1}{\prod_{i<j}(e_i - e_j)}\Bigg),\end{aligned} \tag{4.13}$$

where additional terms obtained by cyclic permutations of e_2, e_3, e_4 are to be added in the first term on the right. We will eventually compare this formula to the known genus 1 partition function of the monster theory of $k = 1$.

In (4.13), we included a factor of $(de)^{3/2}$ for each operator $\mathcal{E}(e)$ to reflect the fact that $\mathcal{E}$ has dimension 3/2. Thus the correlation "function" $\langle\mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$ is most naturally understood as a 3/2-differential in each variable. This factor is often omitted in writing such formulas, and we will do so as an abbreviation except when the factor is important.

For $k = 2$, the singular part of $\mathcal{E}\cdot\mathcal{E}$ cannot be expressed in terms of the diagonal stress tensor T only, and there is no way to avoid using $\mathrm{Sym}^2\mathcal{V}$. Correspondingly, the chiral algebra for $k = 2$ seems to be unfamiliar, and it is considerably harder to get the analog of (4.13) for $k = 2$. Since we do not have an explicit realization of the algebra analogous to (4.12), we cannot compute the four-point function directly, and instead we will determine it by requiring the appropriate OPE singularities.

The function $\langle\mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$, with $e_1,\dots,e_4\in C_0\cong\mathbb{CP}^1$, must be symmetric in all arguments and must have the appropriate OPE singularities in all channels. In addition, it is highly constrained by invariance under the action of $SL(2,\mathbb{C})$ on C_0. The most general "function," actually a cubic differential in each variable, that is symmetric, $SL(2,\mathbb{C})$-invariant, and consistent with at least the leading OPE singularity, is

$$\begin{aligned}\langle\mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle = &\prod_{i=1}^{4}(de_i)^3\Bigg(\left(\frac{1}{(e_1-e_2)^6(e_3-e_4)^6}+\text{cyclic}\right)\\ &+A\left(\frac{1}{(e_1-e_2)^4(e_3-e_4)^4}\prod_{i=1,2,\,j=3,4}\frac{1}{e_i-e_j}+\text{cyclic}\right)\\ &+B\frac{1}{\prod_{i<j}(e_i-e_j)^2}.\Bigg).\end{aligned} \tag{4.14}$$

Cyclic permutations of e_2, e_3, e_4 are to be added as shown, and the coefficients A and B must be determined to ensure that the OPE singularities are correct.

After a detailed calculation, one finds that the right values are

$$A=\frac{3}{16},\ B=\frac{12k+1}{2^{16}} \tag{4.15}$$

where of course $k = 2$. To get these formulas, one may for instance set $e_1 = -e_2 = e$, and use (4.10) to determine the singular behavior for $e \to 0$. In this way, one expresses the singular behavior of the four-point function in terms of three-point functions $\langle\mathcal{X}(0)\mathcal{E}(e_3)\mathcal{E}(e_4)\rangle$, where $\mathcal{X}$ is one of the descendants of the identity that appear in (4.10).

The coefficient A appears in the coefficient of $1/e^4$ for $e \to 0$, and in determining it, the only important choices of $\mathcal{X}$ are 1 and T. The three-point functions that we need are

$$\begin{aligned} \bigl\langle 1 \cdot \mathcal{E}(z)\mathcal{E}(w)\bigr\rangle &= \frac{1}{(z-w)^6} \\ \bigl\langle T(x)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle &= \frac{3}{(z-x)^2(w-x)^2(z-w)^4}. \end{aligned} \tag{4.16}$$

The first is immediate from the way we have normalized $\mathcal{E}$. The second, since the three fields involved are quasi-primary fields (they transform as primaries under $SL(2,\mathbb{C})$) is determined by $SL(2,\mathbb{C})$ invariance up to a constant multiple. The constant can be determined using the $T \cdot \mathcal{E}$ OPE (which reflects the fact that $\mathcal{E}$ is a primary of dimension $3 = 3k/2$) or the $\mathcal{E} \cdot \mathcal{E}$ OPE.

To determine B, we need three more three-point functions, namely

$$\begin{aligned} \bigl\langle \partial^2 T(0)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle &= \left(\frac{1}{z^2} + \frac{1}{w^2}\right)\frac{18}{z^2 w^2 (z-w)^4} + \frac{24}{z^3 w^3 (z-w)^4} \\ \bigl\langle T \star T(0)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle &= \left(\frac{5}{z^2} + \frac{5}{w^2} - \frac{4}{zw}\right)\frac{3}{z^2 w^2 (z-w)^4} \\ \bigl\langle T^+T^-(0)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle &= \left(\frac{9k^2}{16} + \frac{3k}{64}\right)\frac{1}{z^4 w^4 (z-w)^2}. \end{aligned} \tag{4.17}$$

The first of these results is obtained simply by differentiating the second formula in (4.16).

To get the second, we use the fact that

$$T \star T(0) = \mathrm{Res}_{x=0} \frac{dx\, T(x)}{x} T(0). \tag{4.18}$$

Now consider the expression

$$\Bigl\langle \frac{dx\, T(x)}{x} T(0)\mathcal{E}(z)\mathcal{E}(w)\Bigr\rangle. \tag{4.19}$$

We view this as a differential form on the complex x-plane, with z and w kept fixed. As such, it has poles precisely at $x = 0, z, w$. The residue at $x = 0$ is the desired three-point function $\bigl\langle T \star T(0)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle$. The residues at at $x = z, w$ can be computed in terms of the correlator $\bigl\langle T(0)\mathcal{E}(z)\mathcal{E}(w)\bigr\rangle$ by using the fact that $\mathcal{E}$ is a primary of dimension 3, which determines the singular behavior of the product $T \cdot \mathcal{E}$. The sum of the residues must vanish, and this gives our result.

Finally, to get the last formula in (4.17), we note first that since the operators T^+T^- and $\mathcal{E}$ are quasiprimary fields of dimension 4 and 3, we have

$$\left\langle T^+T^-(x)\mathcal{E}(z)\mathcal{E}(w)\right\rangle = \frac{\theta}{(x-z)^4(x-w)^4(z-w)^2}, \tag{4.20}$$

for some constant θ. We have

$$\lim_{w\to\infty}\frac{\left\langle T^+T^-(x)\mathcal{E}(0)\mathcal{E}(w)\right\rangle}{\left\langle\mathcal{E}(0)\mathcal{E}(w)\right\rangle} = \frac{\theta}{x^4}. \tag{4.21}$$

This is supposed to be computed in the theory $\mathrm{Sym}^2\mathcal{W}$ on the x-plane, with insertions of $\mathcal{E}$ at $0, w$. This is equivalent to a computation in the theory $\mathcal{W}$ on a double cover of the x-plane branched at $0, w$. In the limit $w \to \infty$, the double cover is described by the equation $y^2 = x$, and it is convenient to transfer the computation to the y-plane.

We will now write the stress tensor explicitly as T_{xx} or T_{yy} depending on which local parameter is used in defining it. Allowing for the conformal anomaly, the relation between the stress tensor T_{xx} on the x-plane and the corresponding stress tensor T_{yy} on the y-plane is

$$\left(\frac{\partial y}{\partial x}\right)^2 T_{yy} = T_{xx} - \frac{c}{12}\{y,x\}, \tag{4.22}$$

where $\{y,x\}$ denotes the Schwarzian derivative of these two functions. With $y^2 = x$, we get $\{y,x\} = 3/8x^2$. In the case of interest, $c = 24k$ so

$$T_{xx} = \frac{1}{4x}T_{yy} + \frac{3k}{4x^2}. \tag{4.23}$$

This formula enables us to express T^+ in terms of $T_{yy}|_{y=\sqrt{x}}$ and T^- in terms of $T_{yy}|_{y=-\sqrt{x}}$. The limit in (4.21) becomes

$$\left\langle \left(\frac{1}{4x}T_{yy}|_{y=\sqrt{x}} + \frac{3k}{4x^2}\right)\left(\frac{1}{4x}T_{yy}|_{y=-\sqrt{x}} + \frac{3k}{4x^2}\right)\right\rangle. \tag{4.24}$$

(The idea is that passing to the y-plane eliminates the twist operators $\mathcal{E}$ from the formalism, and replaces $T^\pm(x)$ by a stress tensor evaluated at $y = \pm\sqrt{x}$.) Using the fact that in conformal field theory on the y-plane, $\langle T_{yy}\rangle = 0$ and $\left\langle T_{yy}(y_1)T_{yy}(y_2)\right\rangle = c/2(y_1 - y_2)^4$ with $c = 24k$, the correlation function in (4.24) can be evaluated to give θ/x^4 with $\theta = 9k^2/16 + 3k/64$. This gives the last formula in (4.17).

With the aid of the three-point functions in (4.16) and (4.17) and the operator product formula (4.10), one can verify that the candidate four-point function (4.14) has the right singularities precisely if the coefficients A and B are as claimed in (4.15).

Our derivation has in fact been somewhat redundant, since the formulas for three-point functions can largely be deduced from the formula for the $\mathcal{E}\cdot\mathcal{E}$ operator product, and vice-versa.

4.4. The Genus One Partition Function Revisited

We will now explicitly use our formulas for the four-point function of the twist field $\mathcal{E}$ to recover the partition functions of a $k = 1$ or $k = 2$ extremal CFT on a Riemann surface of genus 1.

First we must describe the necessary formalism more precisely. The partition function $Z_C(e_1, e_2, \ldots, e_{2g+2})$ of any conformal field theory $\mathcal{W}$ on the hyperelliptic Riemann surface C defined by

$$y^2 = \prod_{i=1}^{2g+2} (x - e_i) \tag{4.25}$$

can indeed be expressed in terms of the genus zero twist field correlation function

$$\big\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\ldots\mathcal{E}(e_{2g+2})\big\rangle. \tag{4.26}$$

However, to do so requires an important detail that we have not yet described.

There are at first sight several strange features in relating Z_C to the twist field correlation function:

(1) The dimension of the twist field $\mathcal{E}$ seems to be wrong. The Riemann surface C degenerates as $e_i \to e_j$ for any i and j. As the complex structure of C is invariant under the exchange $e_i \leftrightarrow e_j$, a natural local parameter near the degeneration locus on the moduli space is actually $w = (e_i - e_j)^2$. (For example, this statement can be made precise using eqn. (4.30) below.) One expects Z_C to behave for $e_i \to e_j$ as w^{-k}, since the ground state energy is $-k$. However, as $\mathcal{E}$ has dimension $3k/2$, the behavior of the twist field correlation function as $e_i \to e_j$ is actually $(e_i - e_j)^{-3k} \sim w^{-3k/2}$.

(2) This brings in relief the fact that if k is odd, $\mathcal{E}$ actually has half-integral dimension and so is fermionic. For odd k, the twist field correlation function (4.26) actually changes sign under exchange of e_i and e_j, while Z_C is invariant under this exchange.

(3) The twist field correlation function is most naturally understood as a $3k/2$-differential in each variable. We have written the formulas (4.13) and (4.14) in such a way as to emphasize this. It may not be immediately clear what this means for the partition function Z_C.

The answer to all of these questions has to do with the proper treatment of the conformal anomaly in the holomorphic context. We postpone a full explanation to section

4.5 and for now we merely write down the precise relation between the genus 1 partition function and the twist field correlation function:

$$Z_C(e_1, e_2, \ldots, e_4) = 2^{8k}\Bigl(\prod_{1\le i<j\le 4}(e_i - e_j)^k\Bigr)\bigl\langle\mathcal{E}(e_1)\mathcal{E}(e_2)\ldots\mathcal{E}(e_4)\bigr\rangle. \tag{4.27}$$

The factor of 2^{8k} just reflects the way we have normalized $\mathcal{E}$. The factor of $\prod_{i<j}(e_i - e_j)^k$ resolves problems (1) and (2) above. In section 4.5, we will explain the meaning of that factor and its relation to problem (3) (as well as the generalization to higher genus). For now, we merely mention that in eqn. (4.27), the twist field correlation function should be understood as a function, without the factor of $\prod_i (de_i)^{3k/2}$.

Now we will use the general relation (4.27) along with our previous formulas for the twist field correlation functions to analyze the genus 1 partition function for $k = 1, 2$. We make an $SL(2,\mathbb{C})$ transformation to set $e_4 = \infty$ and to fix

$$e_1 + e_2 + e_3 = 0. \tag{4.28}$$

The genus 1 partition function at $k = 1$ is then, using (4.13),

$$Z_C(k=1) = \frac{2^8}{\prod_{i<j}(e_i - e_j)^2}\bigl((e_1 - e_3)^3(e_2 - e_3)^3 + \text{cyclic}\bigr) + 24. \tag{4.29}$$

We want to express this in terms of $q = \exp(2\pi i\tau)$, the usual parameter on the genus 1 moduli space, and to show explicitly that Z_C has the expected behavior near $q = 0$. The classical formulas are

$$\begin{aligned} e_1 - e_2 &= \theta_3^4(0,\tau) \\ e_3 - e_2 &= \theta_1^4(0,\tau) \\ e_1 - e_3 &= \theta_2^4(0,\tau), \end{aligned} \tag{4.30}$$

where

$$\begin{aligned} \theta_1(0,\tau) &= \sum_{n\in\mathbb{Z}} q^{(n+1/2)^2/2} \\ \theta_2(0,\tau) &= \sum_{n\in\mathbb{Z}} (-1)^n q^{n^2/2} \\ \theta_3(0,\tau) &= \sum_{n\in\mathbb{Z}} q^{n^2/2}. \end{aligned} \tag{4.31}$$

Expanding (4.29) near $q = 0$, one can readily verify the expected behavior $Z_C(k=1) = q^{-1} + \mathcal{O}(q)$. Since Z_C is manifestly modular-invariant, this implies that Z_C must equal

J. This can also be verified directly by comparing (4.29) to the classical formula for the j-function

$$j(q) = 2^5 \frac{\left((e_1-e_2)^2+(e_2-e_3)^2+(e_3-e_1)^2\right)^3}{\prod_{i<j}(e_i-e_j)^2} \tag{4.32}$$

and recalling that $J = j - 744$.

The case $k=2$ can be treated similarly. Our previous formula (4.14) for the twist field four-point function implies, via (4.27), a formula for $Z_C(k=2)$. If we take $e_4 \to \infty$ and $e_1+e_2+e_3=0$, this formula becomes

$$\begin{aligned} Z_C(k=2) = & 2^{16}\left(\frac{(e_1-e_3)^2(e_2-e_3)^2}{(e_1-e_2)^4}+\text{cyclic}\right) \\ & + 3\cdot 2^{12}\left(\frac{(e_1-e_3)(e_2-e_3)}{(e_1-e_2)^2}+\text{cyclic}\right)+(12k+1). \end{aligned} \tag{4.33}$$

One can expand this near $q=0$ and verify the expected behavior $Z_C(k=2) = q^{-2}+1+\mathcal{O}(q)$. Together with modular invariance, this implies that $Z_C(k=2) = J^2 - 393767$, as one can also verify directly using (4.32).

4.5. Role Of The Conformal Anomaly

Naively speaking, the partition "function" of a conformal field theory on a Riemann surface C depends only on the complex or conformal structure of C.

Actually, because of the conformal anomaly, things are more subtle. The usual way to describe the situation is to endow C with a metric h, and to think of the partition function as a function $Z(h)$ that transforms in a certain prescribed way (depending on the central charge) under conformal rescaling $h \to e^{\phi}h$ of the metric.

This description is natural in differential geometry. However, it has a few limitations. One is that it really only gives an adequate framework if the left and right central charges c_L and c_R are equal. A second limitation, more crucial for our purposes, is that it is not well-adapted to holomorphic factorization. In a holomorphic CFT, everything will be much simpler if we can incorporate the conformal anomaly in purely holomorphic terms.

An alternative approach is to think of the partition function not as a function but as a section of a suitable line bundle $\mathcal{R}$ over $\mathcal{M}$, the moduli space of complex Riemann surfaces. This approach is developed in detail in [61]. Holomorphic factorization is incorporated naturally in this approach by saying that in a holomorphic CFT, $\mathcal{R}$ is a holomorphic line bundle over $\mathcal{M}$.

More specifically, $\mathcal{R} \cong \mathcal{L}^{c/2}$, where c is the central charge and $\mathcal{L}$ is a fundamental line bundle over $\mathcal{M}$ that is known as the determinant line bundle. $\mathcal{L}$ is defined as follows. Let p be a point in $\mathcal{M}$ corresponding to a Riemann surface C. Then $\mathcal{L}_p$, the fiber of $\mathcal{L}$ at p, is the top exterior power of the vector space $H^0(C, K_C)$ of holomorphic differentials on C. So if C has genus g and $\omega_1, \ldots, \omega_g$ are holomorphic differentials on C, then the expression $\omega_1 \wedge \omega_2 \wedge \ldots \wedge \omega_g$ defines a vector in $\mathcal{L}_p$. In our problem, $c = 24k$, and the partition function is a section of $\mathcal{L}^{12k}$.

On a hyperelliptic Riemann surface

$$y^2 = \prod_{i=1}^{2g+2} (x - e_i), \tag{4.34}$$

we can make this completely explicit. A basis of the space of holomorphic differentials is given by

$$\frac{dx}{y}, \frac{x\,dx}{y}, \frac{x^2\,dx}{y}, \ldots, \frac{x^{g-1}\,dx}{y}. \tag{4.35}$$

A section of $\mathcal{L}^{12k}$ over the space of hyperelliptic equations is hence an expression of the form

$$\Theta = \left(\frac{dx}{y} \wedge \frac{x\,dx}{y} \wedge \frac{x^2\,dx}{y} \wedge \ldots \wedge \frac{x^{g-1}\,dx}{y} \right)^{12k} F(e_1, \ldots, e_{2g+2}). \tag{4.36}$$

To get a section of $\mathcal{L}^{12k}$ over the moduli space of hyperelliptic Riemann surfaces, Θ must be invariant under the action of $SL(2, \mathbb{C})$ on the space of hyperelliptic equations. This action takes the form

$$\begin{aligned} e_i &\to \frac{ae_i + b}{ce_i + d} \\ x &\to \frac{ax + b}{cx + d} \\ y &\to \frac{y}{(cx + d)^{g+1} \prod_{i=1}^{2g+2} (ce_i + d)^{1/2}}, \end{aligned} \tag{4.37}$$

and the condition for Θ to be invariant is that F transforms by

$$F(e_1, \ldots, e_{2g+2}) \to F(e_1, \ldots, e_{2g+2}) \prod_{i=1}^{2g+2} (ce_i + d)^{-6kg}. \tag{4.38}$$

The twist field correlation function, on the other hand, is of the form

$$\left\langle \mathcal{E}(e_1)\mathcal{E}(e_2) \ldots \mathcal{E}(e_{2g+2}) \right\rangle = \prod_{i=1}^{2g+2} (de_i)^{3k/2} G(e_1, \ldots, e_{2g+2}) \tag{4.39}$$

for some "function" G. $SL(2,\mathbb{C})$ invariance of the correlation function means that G transforms by

$$G(e_1,\dots,e_{2g+2}) \to G(e_1,\dots,e_{2g+2}) \prod_{i=1}^{2g+2}(ce_i+d)^{3k}. \tag{4.40}$$

Comparing (4.38) and (4.40), we see that the two transformation laws are consistent if the relation between F and G is

$$F(e_1,\dots,e_{2g+2}) = A_g\, G(e_1,\dots,e_{2g+2}) \prod_{1\le i<j\le 2g+2}(e_i-e_j)^{3k} \tag{4.41}$$

with some constant A_g. Moreover, this is the most general transformation between F and G that is holomorphic, invariant under permutations of the e_i, and an isomorphism as long as the e_i are all distinct. So it must be correct with some choice of the constant. (The relation between F and G is actually part of a general theory that is briefly indicated below.)

This is not the whole story. The partition function has another important property in genus 1: it can be regarded as an ordinary function, rather than a section of a line bundle. Indeed, the genus 1 partition function can be defined as a trace, $Z(q) = \mathrm{Tr}\, q^{L_0}$, and in this form it manifestly is equal to an ordinary function of q. Yet in the above presentation, the genus 1 partition function is a section of $\mathcal{L}^{12k}$. What reconciles the two points of view is that the line bundle $\mathcal{L}^{12}$ is trivial in genus 1. If we describe a genus 1 Riemann surface C by the hyperelliptic equation $y^2 = \prod_{i=1}^{4}(x-e_i)$, then as explained in [61], $\mathcal{L}^{12}$ is trivialized by the section

$$s = \left(\frac{dx}{y}\right)^{12} \prod_{1\le i<j\le 4}(e_i-e_j)^2. \tag{4.42}$$

The point of this formula is that s is holomorphic, $SL(2,\mathbb{C})$-invariant, invariant under permutations of the e_i, and nonzero as long as the e_i are all distinct (so that C is smooth). A power $\mathcal{L}^n$ of the line bundle $\mathcal{L}$ can be trivialized in this fashion precisely if n is an integer multiple of 12 (and this gives one way to understand the fact that holomorphic factorization is possible only if c is an integer multiple of 24). The genus 1 partition function Z understood as an ordinary function is obtained from the section Θ of $\mathcal{L}^{12k}$ by dividing by s^k. Thus

$$Z = \frac{F}{\prod_{1\le i<j\le 4}(e_i-e_j)^{2k}}. \tag{4.43}$$

Combining this with (4.41), we get the relation (4.27) between the genus 1 partition function, understood as a trace, and the correlation function of twist fields:

$$Z = A_1\Big(\prod_{1\le i<j\le 4}(e_i-e_j)\Big)\big\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\mathcal{E}(e_3)\mathcal{E}(e_4)\big\rangle. \tag{4.44}$$

The numerical value of the constant A_1 depends on how the twist field $\mathcal{E}$ has been normalized.

What about $g>1$? What is the best formula depends on how one prefers to treat the conformal anomaly, and there seems to be in the physics literature no standard recipe for doing so in the holomorphic context. To make contact with the genus 2 formulas of Tuite [54], we may proceed as follows. A genus 2 Riemann surface is always a hyperelliptic curve, $y^2=\prod_{i=1}^6(x-e_i)$. In genus 2, we cannot trivialize the line bundle[20] $\mathcal{L}^{12}$, but we can trivialize $\mathcal{L}^{10}$ by an obvious analog of (4.42):

$$\widetilde{s}=\left(\frac{dx}{y}\wedge\frac{x\,dx}{y}\right)^{10}\prod_{1\le i<j\le 6}(e_i-e_j)^2. \tag{4.45}$$

Hence, instead of regarding the genus 2 partition function as a section Θ of $\mathcal{L}^{12k}$, we can regard it as a section $Z=\Theta/\widetilde{s}^k$ of $\mathcal{L}^{2k}$. If we do this, then the relation between Z and the twist field correlation function is the obvious analog of (4.44):

$$Z = A_2\Big(\prod_{1\le i<j\le 6}(e_i-e_j)\Big)\big\langle \mathcal{E}(e_1)\mathcal{E}(e_2)\dots\mathcal{E}(e_6)\big\rangle. \tag{4.46}$$

It can be shown that, in genus 2, a section of $\mathcal{L}^r$ is the same as a Siegel modular function of weight r. So the formula (4.46) is the one we should use if we wish to regard the genus 2 partition function at $c=24k$ as a Siegel modular function of weight $2k$, as was done in [54]. This probably gives the most intuitively appealing formulas in genus 2. For genus greater than 2, we leave the choice of the most useful formalism to the reader.

We conclude by briefly sketching the theoretical context for the formula (4.41). In doing so, we will not limit ourselves to the case of a hyperelliptic Riemann surface, but will consider an arbitrary base curve C_0 and a two-fold cover $C\to C_0$ branched at points $e_1,\dots,e_s$. Let K_i be the cotangent bundle to C_0 at the point e_i. Let $\mathcal{L}_{C_0}$ and $\mathcal{L}_C$ be the determinant lines of C_0 and C respectively. Then as has been shown by P. Deligne, using results in [62], there is a natural isomorphism $\mathcal{L}_C^8\cong\mathcal{L}_{C_0}^{16}\otimes_i K_{C_i}$. If C_0 is of genus zero, then $\mathcal{L}_{C_0}$ is naturally trivial, and the $3k/2$ power of this isomorphism is the map in (4.41). (The isomorphism has a natural square root, relevant if k is odd.)

---

[20] This fact is consistent with holomorphic factorization at $c=24k$, since the genus 2 partition function cannot be written as a trace and so need not be definable as an ordinary function.

Appendix A. More Extremal Partition Functions

F In section 3.3, we have already described the extremal superconformal partition functions F_{k^*} for $k^* \leq 4$. The purpose of this appendix is to briefly describe these functions, and the corresponding Ramond partition functions H_{k^*}, for $5 \leq k^* \leq 10$.

For $k^* = 5$, we have

$$\begin{aligned} F_5 &= K^5 - 1380K^3 - 10239K^2 + 324871K + 2579929 \\ &= q^{-5/2} + q^{-1} + q^{-1/2} + 1 + 924492q^{1/2} + 185710868q + 11953414737q^{3/2} \\ &\quad + 416513865728q^2 + 9727092442216q^{5/2} + \dots . \end{aligned} \tag{A.1}$$

The corresponding Ramond partition function is

$$H_5 = 143+371027968q+832984743936q^2+340658459983872q^3+56418463131631616q^4+\dots . \tag{A.2}$$

The coefficients in these series cannot be expressed as linear combinations of dimensions of monster representations with reasonably small coefficients. But strangely, at least one difference can be so represented; the difference between the coefficient of q in F_5 and half the corresponding coefficient in H_5 is $196884 = 1 + 196883$. (The coefficients in H_5, except the first one, are even because of supersymmetry, so it is natural to divide by 2.) Also, the first coefficient in H_5, namely 143, is the dimension of the smallest nontrivial representation of a sporadic finite group known as the Suzuki group. Unfortunately, the other coefficients are too large, and the representations of the Suzuki group too small, for it to be easy to get convincing evidence of whether the Suzuki group is a symmetry of supergravity at $k^* = 5$.

For $k^* = 6$, the partition functions are

$$\begin{aligned} F_6 &= K^6 - 1656K^4 - 12287K^3 + 618373K^2 + 6487237K - 12026567 \\ &= q^{-3} + q^{-3/2} + q^{-1} + q^{-1/2} + 1 + 3724378q^{1/2} + 1298410586q \\ &\quad + 127852130050q^{3/2} + 6378693040128q^2 + 204402839559265q^{5/2} + \dots \\ H_6 &= 1 + 2589372416q + 12754796707840q^2 \\ &\quad + 9529193701720064q^3 + 2622934801904828416q^4 + \dots . \end{aligned} \tag{A.3}$$

The fact that the leading coefficient in the Ramond sector is 1 suggests that the model may be invariant under a very large discrete symmetry group. The monster $\mathbb{M}$ does not work, since the first non-trivial NS coefficient 3724378 cannot be expressed nicely in terms of dimensions of monster representations. It turns out that another large sporadic group, the baby monster $\mathbb{B}$, is a better candidate. The dimensions $r_1, \ldots, r_{12}$ of the first 12 irreducible representations of $\mathbb{B}$ are listed in Table 3. In terms of those dimensions, we find that the first two non-trivial NS coefficients in the partition functions can be expressed as follows

$$\begin{aligned} 3724378 &= 7r_1 + 4r_2 + 3r_3 + 3r_4 \\ 1298410586 &= 14r_1 + 16r_2 + 7r_3 + 8r_4 + 4r_6 + 3r_7 + r_8 + 2r_9. \end{aligned} \tag{A.4}$$

These formulas are not unique, as the r_i obey linear relations with small coefficients, for example $r_1 + r_3 + r_5 + r_8 = r_2 + r_4 + r_9$. If we let h_1 denote the first non-trivial coefficient in H_6, then it turns out that $h_1/2 = f_1 - f_{1/2} = 1298410586 - 3724378$. So by subtracting the formulas in (A.4), $h_1/2$ can also be expressed as a positive linear combination of the r_i with fairly small coefficients. If is also true that $h_2/2 = f_2 - f_1 + f_{1/2}$, and so is consistent with $\mathbb{B}$ symmetry if f_2 can be suitably expressed in terms of the r_i.

These results are suggestive of baby monster symmetry at $k^* = 6$, though they are perhaps a little less striking than the evidence presented in section 3.3 for monster symmetry at $k^* = 4$, since the required coefficients are larger and the dimensions involved are smaller. Also, it is harder to continue this analysis beyond the first few coefficients, as many representations of $\mathbb{B}$ come in.

r_1	1
r_2	4371
r_3	96255
r_4	1139374
r_5	9458750
r_6	9550635
r_7	63532485
r_8	347643114
r_9	356054375
r_{10}	1407126890
r_{11}	3214743741
r_{12}	4221380670

Table 3. Presented here from [53] are the dimensions r_i of the i^{th} irreducible representation of the baby monster group $\mathbb{B}$, for $i = 1, \ldots, 12$.

We present the remaining cases with little comment. For $k^* = 7$, we have

$$\begin{aligned} F_7 &= K^7 - 1932K^5 - 14335K^4 + 988051K^3 + 11525041K2 - 75824563K - 840705550 \\ &= q^{-7/2} + q^{-2} + q^{-3/2} + q^{-1} + q^{-1/2} + 2 + 13404883q^{1/2} \\ &\quad + 7740446996q + 1126452195714q^{3/2} + 78170641348884q^2 + \dots . \\ H_7 &= 262 + 15394525184q + 156260255891456q^2 + 2032039505847746 56q^3 + \dots . \end{aligned} \tag{A.5}$$

For $k^* = 8$, the analogous formulas are

$$\begin{aligned} F_8 &= K^8 - 2208K^6 - 16383K^5 + 1433905K^4 + 17693341K^3 - 213343055K^2 \\ &\quad - 3164679732K - 2780845557 \\ &= q^{-4} + q^{-5/2} + q^{-2} + q^{-3/2} + q^{-1} + 2q^{-1/2} + 3 + 44146598q^{1/2} \\ &\quad + 40700662036q + 8516908978515q^{3/2} + \dots \\ H_8 &= -213 + 80651894784q + 1605169778655232q^2 + 349692259738655129 6q^3 + \dots . \end{aligned} \tag{A.6}$$

This is the first case in which the leading coefficient in H is negative, so that it is necessary to add an integer to F and H.

For $k^* = 9$, we find

$$\begin{aligned} F_9 &= K^9 - 2484K^7 - 18431K^6 + 1955935K^5 + 24992137K^4 - 445606619K^3 \\ &\quad - 7443672266K^2 + 1774761946K + 223898812203 \\ &= q^{-9/2} + q^{-3} + q^{-5/2} + q^{-2} + q^{-3/2} + 2q^{-1} + 3q^{-1/2} + 3 + 135149374q^{1/2} \\ &\quad + 193216791918q + 56847816152503q^{3/2} + \dots \\ H_9 &= 453 + 381161021440q + 14282018665201664q^2 + 5051928839165624320 0q^3 + \dots \end{aligned} \tag{A.7}$$

Finally, for $k^* = 10$, we find

$$\begin{aligned} F_{10} &= K^{10} - 2760K^8 - 20479K^7 + 2554141K^6 + 33421429K^5 - 793639831K^4 \\ &\quad - 14145708496K^3 + 30831695165K^2 + 1166011724825K + 2482063616019 \\ &= q^{-5} + q^{-7/2} + q^{-3} + q^{-5/2} + q^{-2} + 2q^{-3/2} + 3q^{-1} + 3q^{-1/2} + 3 \\ &\quad + 389274233q^{1/2} + 842231630010q + 341925580784341q^{3/2} + \dots \\ H_{10} &= -261 + 1652836102144q + 112692628289650688q^2 + 63052056690161400 2176q^3 + \dots . \end{aligned} \tag{A.8}$$

This gives a second example in which the leading coefficient of H is negative, so that it is necessary to add an integer.

References

[1] H. Leutwyler, “A 2+1 Dimensional Model For The Quantum Theory Of Gravity,” Nuovo Cim. **42A** (1966) 159.

[2] E. J. Martinec, “Soluble Systems In Quantum Gravity,” Phys. Rev. **D30** (1984) 1198.

[3] S. Deser, R. Jackiw, and G. ’t Hooft, “Three-Dimensional Einstein Gravity: Dynamics of Flat Space,” Annals Phys. **152** (1984) 220.

[4] S. Carlip, “Conformal Field Theory, $(2+1)$-Dimensional Gravity, and the BTZ Black Hole,” gr-qc/0503022.

[5] A. Achúcarro and P. Townsend, “A Chern-Simons Action for Three-Dimensional anti-De Sitter Supergravity Theories,” Phys. Lett. **B180** (1986) 89.

[6] E. Witten, “(2+1)-Dimensional Gravity as an Exactly Soluble System,” Nucl. Phys. **B311** (1988) 46.

[7] A. Ashtekar, “New Variables For Classical And Quantum Gravity,” Phys. Rev. Lett. **57** (1986) 2244-2247.

[8] J. Teschner, “On The Relation Between Quantum Liouville Theory And The Quantum Teichmuller Spaces,” Int. J. Mod. Phys. **A19S2** (2004) 459-477, hep-th/0303149.

[9] J. Teschner, “From Liouville Theory To The Quantum Gravity Of Riemann Surfaces” (International Congress on Mathematical Physics 2003), hep-th/0308031.

[10] J. Teschner, “A Lecture On The Liouville Vertex Operators,” Int. J. Mod. Phys. **A19S2** (2004) 436-458, hep-th/0303150.

[11] J. Teschner, “An Analog Of A Modular Functor From Quantized Teichmuller Theory,” math/0510174.

[12] M. Bañados, C. Teitelboim, and J. Zanelli, “The Black Hole In Three-Dimensional Spacetime,” Phys. Rev. Lett. **69** (1992) 1849-1851, hep-th/9204099.

[13] A. Strominger, “Black Hole Entropy From Near Horizon Microstates,” JHEP 9802:009 (1998), hep-th/9712251.

[14] D. Birmingham, I. Sachs, and S. Sen, “Entropy Of Three-Dimensional Black Holes In String Theory,” Phys. Lett. **B424** (1998) 275-280, hep-th/9801019.

[15] J. Maldacena, “The Large N Limit Of Superconformal Field Theories And Supergravity,” Adv. Theor. Math. Phys. **2** (1998) 231-252, hep-th/9711200.

[16] C. G. Callan, Jr., S. B. Giddings, J. A. Harvey, and A. Strominger, “Evanescent Black Holes,” Phys. Rev. **D45** (1992) 1005-1009, hep-th/9111056.

[17] N. Goheer, M. Kleban, and L. Susskind, “The Trouble With de Sitter Space,” JHEP 0307:066 (2003), hep-th/0212209.

[18] E. Witten, “Quantum Gravity In de Sitter Space,” hep-th/0106109.

[19] A. Strominger, “The dS/CFT Correspondence,” JHEP 0110:034 (2001), hep-th/0106113.

[20] S. Kachru, R. Kallosh, A. Linde, and S. Trivedi, “de Sitter Vacua In String Theory,” Phys. Rev. D68:046005,2003, hep-th/0301240.

[21] J. D. Brown and M. Henneaux, "Central Charges In The Canonical Realization Of Asymptotical Symmetries: An Example From Three-Dimensional Gravity," Commun. Math. Phys. **104** (1986) 207-226.

[22] A. B. Zamolodchikov, "Irreversibility of the Flux of the Renormalization Group in a 2D Field Theory," JETP Lett. **43** (1986) 430.

[23] A. N. Schellekens, "Meromorphic $c = 24$ Conformal Field Theories," Commun. Math. Phys. **153** (1993)159-186, hep-th/9205072.

[24] I. B. Frenkel, J. Lepowsky, and A. Meurman, "A Natural Representation of the Fischer-Griess Monster With the Modular Function J As Character," Proc. Natl. Acad. Sci. USA **81** (1984) 3256-3260.

[25] I. B. Frenkel, J. Lepowsky, and A. Meurman, *Vertex Operator Algebras And The Monster* (Academic Press, Boston, 1988).

[26] L. Dixon, P. Ginsparg, and J. Harvey, "Beauty And The Beast: Superconformal Symmetry In A Monster Module," Commun. Math. Phys. **119** (1988) 221-241.

[27] L. Dolan, P. Goddard, and P. Montague, "Conformal Field Theory, Triality, and the Monster Group," Phys. Lett. **B236** (1990) 165-172.

[28] R. Borcherds, "Automorphic Forms And Lie Algebras," in *Current Developments In Mathematics* (International Press, Boston, 1997).

[29] J. McKay, "The Essentials of Monstrous Moonshine," *Groups and Combinatorics - in Memory of M. Suzuki*, Adv. Studies in Pure Math. **32** (2001) 347-343.

[30] T. Gannon, "Moonshine Beyond The Monster: The Bridge Connecting Algebra, Modular Forms, and Physics" (Cambridge University Press, 2006).

[31] G. Höhn, "Selbstduale Vertexoperatorsuperalgebren und das Babymonster," Ph.D. thesis (Bonn 1995), Bonner Mathematische Schriften 286 (1996), 1-85, arXiv:0706.0236.

[32] G. Höhn, "Conformal Designs based on Vertex Operator Algebras," arXiv:math/0701626.

[33] M. Jankiewicz and T. Kephart, "Modular Invariants And Fischer-Griess Monster," arXiv:math-ph/0608001.

[34] R. Dijkgraaf, J. Maldacena, G. Moore, and E. Verlinde, "A Black Hole Farey Tail," hep-th/0005003.

[35] J. F. Duncan, "Super-Moonshine For Conway's Largest Sporadic Group," arXiv:math/0502267.

[36] S. Deser, R. Jackiw, and S. Templeton, "Topologically Massive Gauge Theories," Annals Phys. **140** (1982) 372-411.

[37] G. Moore and N. Seiberg, "Taming The Conformal Zoo," Phys. Lett. **B220** (1989) 422.

[38] E. Witten,"$SL(2,\mathbb{Z})$ Action On Three-Dimensional Conformal Field Theories With Abelian Symmetry," hep-th/0307041.

[39] S. Carlip "Quantum Gravity In 2+1 Dimensions: The Case Of A Closed Universe," Living Rev. Rel. **8** (2005) 1, gr-qc/0409039.

[40] J. Hartle and S. B. Hawking, "Wavefunction Of The Universe," Phys. Rev. **D28** (1983) 2960.

[41] D. Friedan and S. Shenker, "The Analytic Geometry Of Two-Dimensional Conformal Field Theory," Nucl. Phys. **B281** (1987) 509.

[42] S. Gukov, E. Martinec, G. W. Moore, and A. Strominger, "An Index for 2-D Field Theories with Large $\mathcal{N} = 4$ Superconformal Symmetry," hep-th/0404023.

[43] J. Thompson, "Some Numerology Between the Fischer-Griess Monster and the Elliptic Modular Function," Bull. London Math. Soc. 11 (1979), 352-353.

[44] J. H. Conway and S. P. Norton, "Monstrous Moonshine," Bull. London Math. Soc. **11** (1979) 308-339.

[45] N. Koblitz, *Introduction To Elliptic Curves And Modular Forms* (Springer-Verlag, 1984).

[46] A. Maloney and E. Witten, to appear.

[47] D. Birmingham and S. Sen, "Exact Black Hole Entropy Bound In Conformal Field Theory," hep-th/0008051.

[48] L. Dixon, J. A. Harvey, C. Vafa, and E. Witten, "Strings On Orbifolds," Nucl. Phys. **B261** (1985) 678-686.

[49] M. P. Tuite, "On The Relation Between Generalized Moonshine And The Uniqueness Of The Moonshine Module," Commun. Math. Phys. **166** (1995) 495-532, hep-th/9305057.

[50] J. Conway and N. A. Sloane, *Sphere Packings, Lattices, and Groups* (Grundlehren der Mathematischen Wissenschaften, 1998).

[51] G. Nebe, "Some Cyclo-Quaternionic Lattices," J. Algebra **199** (1998) 472-298.

[52] F. Gliozzi, D. I. Olive, and J. Scherk, "Supersymmetry, Supergravity Theories, And The Dual Spinor Model," Nucl. Phys. **B122** (1977) 253-290.

[53] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas Of Finite Groups* (Oxford University Press, 1985).

[54] M. P. Tuite, "Genus Two Meromorphic Conformal Field Theory," math.qa/9910136.

[55] G. Mason and M. P. Tuite, "On Genus Two Riemann Surfaces Formed From Sewn Tori," Commun. Math. Phys. **270** (2007) 587-634, math.qa/0603088.

[56] Al. B. Zamolodchikov, "Correlation Functions Of Spin Operators In The Ashkin-Teller Model And The Scalar Field On A Hyperelliptic Surface," Soviet Phys. JETP **63** (1986) 1061-1066.

[57] V. G. Knizhnik, "Analytic Fields On Riemann Surfaces," Commun. Math. Phys. **112** (1987) 567.

[58] L. Dixon, D. Friedan, E. Martinec, and S. Shenker, "The Conformal Field Theory Of Orbifolds," Nucl. Phys. **B282** (1987) 13-73.

[59] S. Hamidi and C. Vafa, "Interactions On Orbifolds," Nucl. Phys. **B279** (1987) 465.

[60] A. B. Zamolodchikov, “Infinite Additional Symmetries In Two-Dimensional Conformal Quantum Field Theory,” Theor. Math. Phys. **65** (1985) 1205-1213.

[61] G. Segal, “The Definition Of Conformal Field Theory,” in U. Tillmann, ed. *Topology Geometry, and Quantum Field Theory* (Cambridge University Press, 2004).

[62] P. Deligne, “Le Déterminant de La Cohomologie,” in *Current Trends In Arithmetical Algebraic Geometry*, Contemp. Math. **67** (Amer. Math. Soc., Providence, 1987).

arXiv:1002.3857v2 [hep-th] 23 Feb 2010

Discrete Information from CHL Black Holes

Ashoke Sen

Harish-Chandra Research Institute
Chhatnag Road, Jhusi, Allahabad 211019, India
and
LPTHE, Universite Pierre et Marie Curie, Paris 6
4 Place Jussieu, 75252 Paris Cedex 05, France

E-mail: sen@mri.ernet.in, ashokesen1999@gmail.com

Abstract

AdS_2/CFT_1 correspondence predicts that the logarithm of a $\mathbb{Z}_N$ twisted index over states carrying a fixed set of charges grows as $1/N$ times the entropy of the black hole carrying the same set of charges. In this paper we verify this explicitly by calculating the microscopic $\mathbb{Z}_N$ twisted index for a class of states in the CHL models. This demonstrates that black holes carry more information about the microstates than just the total degeneracy.

Contents

1 Introduction and Summary 2

2 The counting 6

3 Asymptotic Growth 11

4 Conclusion 12

A Explicit computation of $c_b^{(r,s;r',s')}$ 13

B Threshold Integral Representation of $\widetilde{\Phi}$ 14

C Zeroes of $\widetilde{\Phi}$ 19

1 Introduction and Summary

CHL models [1, 2] in four dimensions with $\mathcal{N} = 4$ supersymmetry have proved to be a rich arena for studying the physics of black holes [3–9]. On the one hand they have as much supersymmetry and hence as much control as the toroidally compactified heterotic string theory. On the other hand they have different effective actions beyond the supergravity approximation and hence make different predictions for the entropy of BPS black holes beyond the leading order result of [10, 11]. Thus they provide us with more data points at which we can compare the macroscopic and microscopic predictions for the black hole entropy. This comparison has been remarkably successful at the level of four derivative corrections to the effective action, reproducing complicated non-trivial functional dependence of the entropy on the charges on both sides.[1] Indeed, most of the results on black holes in heterotic string theory on T^6 [6, 16–24] have now been generalized to the case of CHL models.

In this paper we shall make use of the CHL model to explore another aspect of black holes. Based on AdS_2/CFT_1 correspondence [25, 26] it was argued in [27] that if a theory has a $\mathbb{Z}_N$ symmetry that cannot be regarded as part of a $U(1)$ gauge transformation, and if we pick a black hole carrying $U(1)$ charges which are invariant under this $\mathbb{Z}_N$ transformation, then the logarithm of the trace of the $\mathbb{Z}_N$ generator over the microstates of the black hole grows as $1/N$ times the entropy of the black hole.[2] This can be made more concrete for BPS black holes in supersymmetric string theories by working with protected helicity trace index. In the context of $\mathcal{N}=4$ supersymmetric string theories in four dimensions the relevant twisted index is the 6th helicity trace index [27,30,31]:

---

[1] We should add a note of caution that this comparison requires us to make assumption of certain non-renormalization results which have not been proven. In particular it assumes that at the level of four derivative terms the Gauss-Bonnet terms (or their supersymmetric completion given in [12–15]) in the action are sufficient to calculate the correction to the black hole entropy. The analysis in this paper does not require us to make any such assumption.

$$B_6^g(\vec{q}) = \frac{1}{6!}\, Tr_{\vec{q}} \left\{(-1)^{2h}(2h)^6\, g\right\} \,, \tag{1.1}$$

where the trace is taken over all states carrying a fixed set of charges $\vec{q}$, h is the third component of the angular momentum of the state in its rest frame, and g is the generator of a $\mathbb{Z}_N$ symmetry which leaves $\vec{q}$ invariant. This index receives contribution from 1/4 BPS states in this theory which describe large black holes with near horizon $AdS_2 \times S^2$ geometry. In this case the analysis of [27] applies and tells us that

$$|B_6^g(\vec{q})| \sim \exp[S_{BH}(\vec{q})/N]\,, \tag{1.2}$$

where $S_{BH}(\vec{q})$ is the entropy of an extremal black hole carrying charge $\vec{q}$. We shall not review the arguments of [27] here; but the central idea is that in computing the contribution to (1.2) from the horizon of the black hole the leading saddle point corresponding to the $AdS_2 \times S^2$ near horizon geometry does not contribute. However a $\mathbb{Z}_N$ orbifold of $AdS_2 \times S^2$ [26,32,33], whose asymptotic geometry coincides with that of the original near horizon geometry of the black hole, contributes and gives the answer $\exp[S_{BH}/N]$ in the semiclassical limit.

While (1.2) follows almost trivially from the AdS_2/CFT_1 correspondence, it is quite striking from the point of view of the microscopic theory. For large black holes the right hand side of (1.2) is much smaller than the untwisted helicity trace index carrying the same charges, since the latter is given by $\exp[S_{BH}(\vec{q})]$. What this tells us is that in a given charge sector the microstates of different g eigenvalues come in almost equal numbers so that the sum weighted by g is much smaller than the total number of states. This was explicitly verified in [27] by deriving the microscopic formula for this twisted index in toroidally compactified heterotic and type II string theories and then studying their asymptotic behaviour.[3]

---

[2] For a $\mathbb{Z}_N$ group that can be regarded as a subgroup of a spontaneously broken $U(1)$ gauge group, the possibility of hair modes containing information about the $\mathbb{Z}_N$ quantum numbers was explored in [28,29]. In contrast the $\mathbb{Z}_N$ groups we discuss here cannot be regarded as a subgroup of a spontaneously broken $U(1)$ symmetry. Also our goal here is quite different from the one of [28,29].

[3] Even though type II string theory on T^6 has $\mathcal{N}=8$ supersymmetry, only an $\mathcal{N}=4$ subgroup of this commutes with g. Thus effectively we can analyze it in the same way as in an $\mathcal{N}=4$ supersymmetric theory.

Given the unusual nature of this macroscopic prediction, it is important to test this in as many examples as possible. In this paper we shall verify this in the context of CHL models. The construction of the CHL models that we shall analyze proceeds as follows. We begin with type IIB string theory on $\mathcal{M} \times S^1 \times \widetilde{S}^1$ where $\mathcal{M}$ is either K3 or T^4 and go to a special subspace of the moduli space of $\mathcal{M}$ where the theory has a geometric $\mathbb{Z}_M \times \mathbb{Z}_N$ symmetry that commutes with 16 supersymmetry generators of the theory. An extensive list of possible symmetries of this type can be found in [34, 35]. Note also that a Z_{MN} group with M and N relatively prime can be considered as a $Z_M \times Z_N$ group for the purpose of our analysis. Let us denote by g_M and g_N the generators of $\mathbb{Z}_M$ and $\mathbb{Z}_N$ respectively. We now take an orbifold of this theory by a symmetry that involves $1/M$ unit of translation along the circle S^1 accompanied by the transformation g_M. This gives a theory with $\mathcal{N} = 4$ supersymmetry in four dimensions and the $\mathbb{Z}_N$ group generated by g_N is a symmetry of this theory. We now consider a g_N invariant charge vector $\vec{q}$ in this theory and define the index

$$d(\vec{q}) = -\frac{1}{6!}\, Tr_{\vec{q}} \left(e^{2\pi i h} (2h)^6 g_N \right) , \tag{1.3}$$

where the trace is taken over all states carrying the charge $\vec{q}$. Eq.(1.2) now translates to

$$|d(\vec{q})| \sim \exp[S_{BH}(\vec{q})/N] , \tag{1.4}$$

for large charges. Our goal will be to verify this by explicit computation of $d(\vec{q})$ in the microscopic theory.

Since the explicit counting of states involves technical details, we shall take this opportunity to summarize the results of our analysis. We use a convention in which the coordinate radius of the original circle S^1 before orbifolding is $2\pi M$ so that the orbifold action involves translation by 2π along S^1 accompanied by g_M. In this convention the minimum amount of momentum along S^1 is $1/M$. We focus on states carrying one unit of KK monopole charge associated with the circle $\widetilde{S}^1$, one unit of D5-brane charge wrapped on $\mathcal{M} \times S^1$, Q_1 units of D1-brane charge wrapped on S^1, left-moving momentum n/M along S^1 and J units of momentum along $\widetilde{S}^1$, and define

$$Q^2 = 2n/M, \qquad P^2 = 2Q_1, \qquad Q.P = J . \tag{1.5}$$

In this case our result of $d(\vec{q})$ is given by

$$d(\vec{q}) = \frac{1}{M}(-1)^{Q.P+1} \int_{\mathcal{C}} d\rho d\sigma dv e^{-\pi i (M\rho Q^2 + \sigma P^2/M + 2vQ.P)} \frac{1}{\widetilde{\Phi}(\rho, \sigma, v)} . \tag{1.6}$$

Here $\mathcal{C}$ is a three real dimensional subspace of the three complex dimensional space labelled by $(\rho=\rho_1+i\rho_2, \sigma=\sigma_1+i\sigma_2, v=v_1+iv_2)$ given by

$$\begin{aligned} &\rho_2 = M_1, \qquad \sigma_2 = M_2, \qquad v_2 = M_3, \\ &0 \le \rho_1 \le 1, \qquad 0 \le \sigma_1 \le M, \qquad 0 \le v_1 \le 1\,, \end{aligned} \tag{1.7}$$

M_1, M_2, M_3 being large but fixed positive numbers satisfying

$$M_1 M_2 > M_3^2\,. \tag{1.8}$$

The function $\widetilde{\Phi}(\rho,\sigma,v)$ is a modular form of a subgroup of $Sp(2,\ \mathbb{Z})$, given by

$$\begin{aligned} \widetilde{\Phi}(\rho,\sigma,v) &= e^{2\pi i(\widetilde{\alpha}\rho+\widetilde{\gamma}\sigma+\widetilde{\beta}v)} \prod_{b=0}^{1}\prod_{r=0}^{N-1}\prod_{r'=0}^{M-1} \prod_{\substack{k\in\mathbb{Z}+\frac{r'}{M}, l\in\mathbb{Z}, j\in 2\mathbb{Z}+b\\ k,l\ge 0, j<0\, for\, k=l=0}} \left[1-e^{2\pi i r/N}\, e^{2\pi i(k\sigma+l\rho+jv)}\right]^{a} \\ a &\equiv \sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i(s'l/M+rs/N)} c_b^{(0,s;r',s')}(4kl-j^2)\,, \end{aligned} \tag{1.9}$$

where the coefficients $c_b^{(r,s;r',s')}$ are defined via the equation:

$$\begin{aligned} &\sum_{b=0}^{1}\sum_{j\in 2\mathbb{Z}+b, n\in\mathbb{Z}/MN} c_b^{(r,s;r',s')}(4n-j^2)e^{2\pi i(n\tau+jz)} \\ &= \frac{1}{MN} Tr_{RR;g_M^{r'}g_N^{r}}\left(g_M^{s'}g_N^{s}(-1)^{J_L+J_R}e^{2\pi i(\tau L_0-\bar{\tau}\bar{L}_0)}e^{2\pi i J_L z}\right)\,. \end{aligned} \tag{1.10}$$

The trace is taken over all the $g_M^{r'}g_N^{r}$ twisted RR sector states in the (4,4) superconformal CFT$_2$ with target space $\mathcal{M}$. L_0 and $\bar{L}_0$ are the left and right-moving Virasoro generators and $J_L/2$ and $J_R/2$ are the generators of the $U(1)_L\times U(1)_R$ subgroup of the $SU(2)_L\times SU(2)_R$ R-symmetry group of this CFT$_2$. An algorithm for explicitly computing the right hand side of (1.10) has been outlined in appendix A. The coefficients $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\gamma}$ are given by

$$\begin{aligned} \widetilde{\alpha} &= \frac{1}{24M}Q_{0,0} - \frac{1}{2M}\sum_{s'=1}^{M-1} Q_{0,s'}\frac{e^{-2\pi i s'/M}}{(1-e^{-2\pi i s'/M})^2}\,, \\ \widetilde{\beta} &= 1 \\ \widetilde{\gamma} &= \frac{1}{24M}\,\chi(\mathcal{M}) = \frac{1}{24M}\,Q_{0,0}\,, \end{aligned} \tag{1.11}$$

where

$$Q_{r',s'} = MN\left(c_0^{(0,0;r',s')}(0) + 2c_1^{(0,0;r',s')}(-1)\right)\,. \tag{1.12}$$

Eqs. (1.10), (1.11) define all the quantities which appear in the definition of $\widetilde{\Phi}$. The only ambiguity that remains in computing the right hand side of (1.6) is the choice of the integration contour encoded in the choice of (M_1, M_2, M_3). As is well known by now, this ambiguity is related to the phenomenon of wall crossing [36–39]. Different choices of M_1, M_2 and M_3 give the value of $d(\vec{q})$ for different values of the asymptotic moduli. However the ambiguity in the value of $d(\vec{q})$ that it introduces is sufficiently small so as not to affect our analysis, and hence we shall ignore it in our subsequent discussion.

Given the result (1.6) for $d(\vec{q})$ we can find its behaviour for large Q^2, P^2 and $Q.P$ by standard method [3,6,16,17]. The result is that $d(\vec{q})$ behaves as

$$d(\vec{q}) \sim \exp\left[\pi\sqrt{Q^2P^2 - (Q.P)^2}/N\right] . \tag{1.13}$$

Since in this limit a black hole of charge $\vec{q}$ has entropy [10,11]

$$S_{BH}(\vec{q}) \simeq \pi\sqrt{Q^2P^2 - (Q.P)^2} , \tag{1.14}$$

we see that the microscopic result (1.13) is in perfect agreement with the macroscopic prediction (1.4).

Finally we would like to remark that even though we have presented our analysis for the index $Tr((-1)^{2h}(2h)^6 g_N)$, we can repeat the analysis with g_N replaced by $(g_N)^b$ for any integer b. In this case the role of N is played by the order of $(g_N)^b$, and in all the formulæ we simply have to replace g_N by $(g_N)^b$. This in turn allows us to compute the index $Tr((-1)^{2h}(2h)^6)$ for states carrying a definite g_N eigenvalue $e^{2\pi i a/N}$ using the combination

$$\frac{1}{N}\sum_{b=0}^{N-1} e^{-2\pi i a b/N}\, Tr((-1)^{2h}(2h)^6(g_N)^b) . \tag{1.15}$$

Thus our result can also be interpreted as the agreement between the macroscopic and the microscopic results for the helicity trace index over states carrying a definite g_N charge.

2 The counting

The counting of states of the D1-D5-KK monopole system proceeds as in [6–8,40]. We take the circle S^1 to be large compared to the size of $\mathcal{M}$ and regard the world-volume theory as a 1+1 dimensional field theory living on S^1. Denoting by $d(Q_1, n, J)$ the twisted index (1.3) of states carrying charge labeled by (Q_1, n, J) in the convention of §1, we define

$$Z(\rho, \sigma, v) = \sum_{Q_1, n, J} e^{2\pi i(Q_1\sigma/M + n\rho + vJ)}(-1)^J d(Q_1, n, J) . \tag{2.1}$$

Proving (1.6) is now equivalent to proving that $Z = -1/\widetilde{\Phi}$.

The twisted partition function Z is given by the product of three twisted partition functions, – that of the excitations living on the KK monopole, that of the dynamics describing the overall motion of the D1-D5 system in the background of the KK monopole and that of the motion of the D1-brane along the D5-brane. For each system we must keep the right-movers in their ground state and excite the left-movers in order to preserve supersymmetry.[4] Since the fermion zero modes associated with the broken supersymmetries are automatically removed while computing the helicity trace (which is B_6 in this case since the system breaks 12 of the 16 supersymmetries), we shall ignore their contribution during the rest of our analysis.

We begin by analyzing the partition function of the KK monopole. The massless bosonic modes on the world-volume of the KK monopole arise from the motion along the three transverse directions and the components of the p-form fields along the product of the harmonic $(p-2)$-forms of $\mathcal{M}$ and the harmonic 2-form of the Taub-NUT space. The massless fermions are the goldstinos associated with the supersymmetries broken by the Kaluza-Klein monopole. In general one can show that the left-moving bosons and fermions are in one to one correspondence with the even and odd degree harmonic forms on $\mathcal{M}$ [40]. Furthermore their (g_M, g_N) quantum numbers are also given by the (g_M, g_N) eigenvalues of the harmonic forms on $\mathcal{M}$. Since the harmonic (p,q) forms on $\mathcal{M}$ are in one to one correspondence with the RR sector ground states in the supersymmetric σ-model with target space $\mathcal{M}$ carrying $L_0 = \bar{L}_0 = 0$, $J_L = (p-1)$, $J_R = (q-1)$, it follows from (1.10) that the number of left-moving bosons minus the number of left-moving fermions on the KK monopole world-volume, carrying g_N quantum number $e^{2\pi i r/N}$ and g_M quantum number $e^{2\pi i k'/M}$, is given by [8, 40]

$$\begin{aligned}
&\frac{1}{MN}\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i r s/N} e^{-2\pi i k' s'/M} Tr_{RR;I}\left[(-1)^{J_L+J_R}\, g_M^{s'} g_N^{s}\, \delta_{L_0,0}\, \delta_{\bar{L}_0,0}\right] \\
= \; &\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i r s/N} e^{-2\pi i k' s'/M}\left(c_0^{(0,s;0,s')}(0) + 2c_1^{(0,s;0,s')}(-1)\right) . \qquad (2.2)
\end{aligned}$$

In arriving at (2.2) we have used the fact that $c_b^{(r,s;r',s')}(u) = 0$ for $u < -1$. Now consider a mode carrying g_M eigenvalue $e^{2\pi i k'/M}$ and left-moving momentum l/M along S^1. The requirement of invariance under the simultaneous action of g_M and 2π translation along S^1 gives us the requirement that $l = k'$ mod M. Furthermore the contribution to the twisted index from these states will be weighted by the g_N eigenvalue $e^{2\pi i r/N}$. Thus the net contribution to the partition function from these modes is given by

---

[4] Here and elsewhere left-moving modes will refer to modes carrying momentum along the negative S^1 direction. Thus left-moving momentum n will indicate momentum $-n$ along S^1.

$$Z_{KK} = e^{-2\pi i\widetilde{\alpha}\rho} \prod_{r=0}^{N-1} \prod_{l=1}^{\infty} \left(1 - e^{2\pi i r/N} e^{2\pi i l\rho}\right)^{-\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i rs/N} e^{-2\pi i l s'/M} \left(c_0^{(0,s;0,s')}(0) + 2c_1^{(0,s;0,s')}(-1)\right)} . \tag{2.3}$$

Here the term $e^{-2\pi i\widetilde{\alpha}\rho}$ reflects the effect of the momentum carried by the ground state of the Kaluza-Klein monopole. The analysis of [8, 40] gives

$$\begin{aligned} \widetilde{\alpha} &= \frac{1}{24M} Q_{0,0} - \frac{1}{2M} \sum_{s'=1}^{M-1} Q_{0,s'} \frac{e^{-2\pi i s'/M}}{(1 - e^{-2\pi i s'/M})^2} , \\ & Q_{r',s'} \equiv MN \left(c_0^{(0,0;r',s')}(0) + 2c_1^{(0,0;r',s')}(-1)\right) . \end{aligned} \tag{2.4}$$

Next we turn to the dynamics of the overall motion of the D1-D5 system in the KK monopole background. The dynamics in the transverse direction is independent of whether we are working with $K3$ or T^4. Furthermore these modes do not carry any g_N or g_M quantum numbers; thus the contribution from these modes to the partition function is universal. The result is [8, 40]

$$\begin{aligned} & -e^{-2\pi i v} \left(1 - e^{-2\pi i v}\right)^{-2} \\ & \prod_{\substack{l \in M\mathbb{Z} \\ l>0}} \{(1 - e^{2\pi i l\rho})^4 (1 - e^{2\pi i l\rho + 2\pi i v})^{-2} (1 - e^{2\pi i l\rho - 2\pi i v})^{-2}\} . \end{aligned} \tag{2.5}$$

The first line represents the contribution from the zero mode dynamics that binds the D1-D5 system to the KK monopole [6, 41, 42], and the second line represents the contribution from the oscillators. The last two terms in the second line of (2.5) represent the contribution from the four left-moving bosonic modes representing transverse oscillation of the D1-D5 system whereas the first factor in the same line represents contribution from the left-moving fermionic modes.[5] In arriving at (2.5) one needs to use the fact that in the presence of the KK monopole background, the momentum along $\widetilde{S}^1$ appears as the angular momentum $2J_L$ for the D1-D5 system where J_L is the generator of the $U(1)_L \subset SU(2)_L$ subgroup of the rotation group in transverse space [43]. The v dependence of (2.5) then follows from the fact that the bosonic modes, transforming as a vector of the transverse rotation group $SU(2)_L \times SU(2)_R$, carry $J_L = \pm 1$ while the fermionic modes are neutral under $U(1)_L$ as a consequence of footnote 5.

---

[5] These left-moving bosonic and fermionic modes, as well as those which contribute to (2.7), are paired by the unbroken superysymmetry transformations on the D1-D5 world volume in flat space-time which commute with $\mathbb{Z}_M \times \mathbb{Z}_N$, are charged under the $SU(2)_L$ subgroup of the transverse rotation group, and act on the left-movers. Eventually when we place this system in the background of KK monopole this supersymmetry is broken since in the full system there is no supersymmetry acting on the left-movers. However this is still useful for determining the quantum numbers of the fermions from the known quantum numbers of the bosonic modes [8, 40].

For $\mathcal{M} = T^4$ we also have four additional bosonic modes arising from the Wilson lines on the D5-brane along T^4 and four additional fermionic modes. In order to find the contribution to the partition function from these modes we need to know the action of g_M and g_N on these modes. If z_1 and z_2 denote the complex coordinates on T^4 then in order to preserve supersymmetry both g_M and g_N must act as equal and opposite rotation of z_1 and z_2, possibly accompanied by shifts. We shall assume for definiteness that g_M and g_N induce respectively $2\pi/M$ and $2\pi/N$ rotations on these coordinates:

$$\begin{aligned} g_M &: \quad (dz_1, dz_2) \to \left(e^{2\pi i/M} dz_1, e^{-2\pi i/M} dz_2\right) , \\ g_N &: \quad (dz_1, dz_2) \to \left(e^{2\pi i/N} dz_1, e^{-2\pi i/N} dz_2\right) . \end{aligned} \tag{2.6}$$

(2.6) represents the action of g_M and g_N on the Wilson line variables. Furthermore the Wilson lines are neutral under the rotation group in the transverse space and hence carry $J_L = 0$. The result of footnote 5 now tells us that the additional fermionic modes on the D1D5 system, which arise for $\mathcal{M} = T^4$, transform in the same way under g_M and g_N, and carry $J_L = \pm 1$ uncorrelated with their (g_M, g_N) quantum numbers [8, 40]. Thus the contribution from these additional modes to the twisted partition function is given by

$$\begin{aligned} &\prod_{\substack{l \in M\mathbb{Z}+1 \\ l>0}} \left(1 - e^{2\pi i/N} e^{2\pi i l\rho}\right)^{-2} \prod_{\substack{l \in M\mathbb{Z}-1 \\ l>0}} \left(1 - e^{-2\pi i/N} e^{2\pi i l\rho}\right)^{-2} \prod_{\substack{l \in M\mathbb{Z}+1 \\ l>0}} \left(1 - e^{2\pi i/N} e^{2\pi i l\rho + 2\pi i v}\right) \\ &\prod_{\substack{l \in M\mathbb{Z}+1 \\ l>0}} \left(1 - e^{2\pi i/N} e^{2\pi i l\rho - 2\pi i v}\right) \prod_{\substack{l \in M\mathbb{Z}-1 \\ l>0}} \left(1 - e^{-2\pi i/N} e^{2\pi i l\rho + 2\pi i v}\right) \prod_{\substack{l \in M\mathbb{Z}-1 \\ l>0}} \left(1 - e^{-2\pi i/N} e^{2\pi i l\rho - 2\pi i v}\right) . \end{aligned} \tag{2.7}$$

The first two factors come from the bosonic modes and the last four factors arise from the fermionic modes whose contribution have not already been included in (2.5). The only new ingredient in this formula compared to that in [8, 40] is the insertion of the factors of $e^{\pm 2\pi i/N}$, – these arise from the insertion of g_N into the trace.

The product of (2.5) and (2.7) can be written in a compact form using the coefficients $c_1^{(0,s;0,s')}(-1)$. It follows from its definition, and the identification of the RR sector ground states in the SCFT with target space $\mathcal{M}$ carrying $(J_L, J_R) = (p-1, q-1)$ with the harmonic

(p,q) forms on $\mathcal{M}$, that $MNc_1^{(0,s;0,s')}(-1)$ represents trace over the $(0,q)$ forms on $\mathcal{M}$ weighted by $(-1)^q g_N^s g_M^{s'}$ [8, 40]. On $K3$ the only $(0,q)$ forms are $(0,0)$ forms and $(0,2)$ forms both of which are neutral under g_N and g_M, while on T^4 we also have a pair of $(0,1)$ forms dz_1 and dz_2 which we have chosen to carry (g_N, g_M) eigenvalues $(e^{\pm 2\pi i/N}, e^{\pm 2\pi i/M})$. This gives

$$\begin{aligned} c_1^{(0,s;0,s')}(-1) &= \frac{2}{MN} \quad \text{for } \mathcal{M} = K3\,, \\ &= \frac{1}{MN}\left(2 - e^{2\pi i s/N} e^{2\pi i s'/M} - e^{-2\pi i s/N} e^{-2\pi i s'/M}\right) \quad \text{for } \mathcal{M} = T^4\,. \end{aligned} \tag{2.8}$$

Using this we can express the total contribution to the partition function from the overall motion of the D1-D5 system in the Taub-NUT space, given by (2.5) for $\mathcal{M} = K3$ and the product of (2.5) and (2.7) for $\mathcal{M} = T^4$, as

$$\begin{aligned} Z_{CM} &= -e^{-2\pi i v} \prod_{l=1}^{\infty} \prod_{r=0}^{N-1} (1 - e^{2\pi i r/N}\, e^{2\pi i l\rho})^{2\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i l s'/M} e^{-2\pi i r s/N} c_1^{(0,s;0,s')}(-1)} \\ & \prod_{l=1}^{\infty} \prod_{r=0}^{N-1} (1 - e^{2\pi i r/N}\, e^{2\pi i l\rho + 2\pi i v})^{-\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i l s'/M} e^{-2\pi i r s/N} c_1^{(0,s;0,s')}(-1)} \\ & \prod_{l=0}^{\infty} \prod_{r=0}^{N-1} (1 - e^{2\pi i r/N}\, e^{2\pi i l\rho - 2\pi i v})^{-\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i l s'/M} e^{-2\pi i r s/N} c_1^{(0,s;0,s')}(-1)}\,. \end{aligned} \tag{2.9}$$

Note that the $(1 - e^{-2\pi i v})^{-2}$ factor has been absorbed into the $l = 0$ term in the last term.

Finally let us turn to the contribution to the partition function from the motion of the D1-branes along the D5-branes. First we consider a single D1-brane wrapped w times along S^1, and count the number of states $n(w,j,l;r,k')$ of the system carrying left-moving momentum l/M along S^1, g_M eigenvalue $e^{2\pi i k'/M}$, g_N eigenvalue $e^{2\pi i r/N}$ and $\widetilde{S}^1$ momentum j. Since the boundary condition on various fields are twisted by g_M under 2π translation along S^1, the CFT on a D1-brane wrapped w times along S^1 satisfies boundary condition twisted by $(g_M)^w$. Furthermore since the effective length of the D1-brane is now $2\pi w$, a momentum l/M along S^1 will appear as lw/M units of momentum in the CFT living on the D1-brane. It now follows from (1.10) that [8, 40]

$$\begin{aligned} n(w,j,l;r,k') &= \sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i r s/N} e^{-2\pi i k' s'/M} c_b^{(0,s;r',s')}(4lw/M - j^2)\,, \\ & b = j \bmod 2, \quad r' = w \bmod M\,. \end{aligned} \tag{2.10}$$

The requirement that we only keep the modes which are invariant under the transformation g_M accompanied by 2π translation along S^1 forces the constraint $k' = l \bmod M$. It is now straightforward to evaluate the contribution to the g_N twisted partition function from multiple states of this type, carrying different (w,l,j) [44]:

$$Z_{D1D5} = e^{-2\pi i\widetilde{\gamma}\sigma} \prod_{r=0}^{N-1}\prod_{b=0}^{1} \prod_{\substack{w\in\mathbb{Z},l\in\mathbb{Z},j\in 2\mathbb{Z}+b\\ w>0,l\geq 0}} \left(1-e^{2\pi i r/N}e^{2\pi i(\sigma w/M+\rho l+vj)}\right)^{-n(w,j,l;r,l)} . \tag{2.11}$$

where

$$\widetilde{\gamma} = \begin{cases} 1/M \text{ for } \mathcal{M}=K3 \\ 0 \text{ for } \mathcal{M}=T^4 \end{cases} . \tag{2.12}$$

The $e^{-2\pi i\widetilde{\gamma}\sigma}$ in (2.11) accounts for the fact that the actual number of D1-branes required to produce a total D1-brane charge Q_1 in the background of a D5-brane is given by Q_1+1 for $\mathcal{M}=K3$ and Q_1 for $\mathcal{M}=T^4$. Multiplying (2.3), (2.9) and (2.11) we get the total partition function of the system:

$$Z(\rho,\sigma,v) = Z_{KK}\, Z_{CM}\, Z_{D1D5} = -1/\widetilde{\Phi}(\rho,\sigma,v)\,, \tag{2.13}$$

where

$$\begin{aligned} \widetilde{\Phi}(\rho,\sigma,v) &= e^{2\pi i(\widetilde{\alpha}\rho+\widetilde{\gamma}\sigma+\widetilde{\beta}v)} \prod_{b=0}^{1}\prod_{r=0}^{N-1}\prod_{r'=0}^{M-1} \prod_{\substack{k\in\mathbb{Z}+\frac{r'}{M},l\in\mathbb{Z},j\in 2\mathbb{Z}+b\\ k,l\geq 0,j<0\; for\; k=l=0}} \left[1-e^{2\pi i r/N}\, e^{2\pi i(k\sigma+l\rho+jv)}\right]^{a} \\ a &\equiv \sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} e^{-2\pi i(s'l/M+rs/N)} c_b^{(0,s;r',s')}(4kl-j^2)\,, \end{aligned} \tag{2.14}$$

with $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\gamma}$ defined in (1.11). Note that the $k=0$ term in this product gives the result for $Z_{KK}Z_{CM}$.

3 Asymptotic Growth

We now study the growth of the index for large Q^2, P^2 and $Q.P$. This can be done by standard procedure described in [3, 6, 16, 17]. We deform the three dimensional contour of integration over (ρ,σ,v) to small imaginary values of (ρ,σ,v). During this deformation we pick up contribution from the residues at various poles, given by the zeroes of $\widetilde{\Phi}$, which give the leading contribution to the index, – the contribution from the final contour can be shown to be subleading compared to the contribution from the residues at the poles [40]. Thus we need to first determine the location of the zeroes of $\widetilde{\Phi}$. This has been done in appendix C where it is shown that $\widetilde{\Phi}$ has double zeroes on the subspaces:

$$n_2(\rho\sigma - v^2) - m_1\rho + n_1\sigma + m_2 + jv = 0\,, \tag{3.1}$$

for values of (m_1, n_1, m_2, n_2, j) satisfying

$$\begin{aligned} & m_1 n_1 + m_2 n_2 + \frac{j^2}{4} = \frac{1}{4}\,, \\ & m_1 \in M\,\mathbb{Z}, \quad m_2 \in \mathbb{Z}, \quad n_2 \in N\,\mathbb{Z}, \quad n_1 \in \mathbb{Z}, \quad j \in 2\,\mathbb{Z}+1\,. \end{aligned} \tag{3.2}$$

Now the analysis of [3,6,16,17] tells us that for large Q^2, P^2, $Q.P$ the contribution from the residue at the pole (3.1) of $1/\widetilde{\Phi}$ grows as

$$\exp\left(\pi\sqrt{Q^2P^2 - (Q.P)^2}/|n_2|\right) \qquad \text{for } |n_2| > 0\,. \tag{3.3}$$

On the other hand the poles at $n_2 = 0$ are responsible for wall crossing and their contribution grows much slower than (3.3) [36–39]. Thus the leading contribution comes from the pole at (3.1) for the minimum non-zero value of $|n_2|$. Eq.(3.2) shows that this is N. Thus the index grows as

$$\exp\left(\pi\sqrt{Q^2P^2 - (Q.P)^2}/N\right)\,. \tag{3.4}$$

Since for this charge the black hole entropy S_{BH} is given by $\pi\sqrt{Q^2P^2 - (Q.P)^2}$ [10,11], (3.4) is in precise agreement with the macroscopic prediction (1.2).

4 Conclusion

It is widely believed that since string theory provides us with a consistent quantum theory of gravity, black holes in string theory do not lead to a loss of information. If so, the black hole must represent an ensemble of microstates and the black hole entropy must have an interpretation as the logarithm of the degeneracy of microstates. Furthermore quantum string theory around a black hole background must contain all possible information about the microstates. It is therefore important to learn how we can extract information about the black hole microstates by studying quantum string theory around the black hole background.

The results of [27] and this paper provide a small step in this direction. In these papers we discuss how to extract information about one specific feature of the black hole microstates, namely distribution of the $\mathbb{Z}_N$ charges among the microstates. Quantum string theory around the near horizon background leads to a specific algorithm for extracting this information. Our analysis shows that in the limit of large charges the results of the macroscopic analysis are in exact agreement with the microscopic results in a wide class of models where the latter is computable. While using the rules of AdS_2/CFT_1 correspondence we can in principle compute the ensemble average of more general operators on the black hole side, in the absence of non-renormalization results it is not clear how we might compare this with the microscopic results.

Acknowledgment: I wish to thank Nabamita Banerjee, Atish Dabholkar, Joao Gomes and Sameer Murthy for useful discussions. This work was supported in part by the JC Bose fellowship of the Department of Science and Technology, India and by the Blaise Pascal Chair, France.

Note added: I have been informed by Suresh Govandarajan that the modular forms of subgroups of $Sp(2, \mathbb{Z})$ which appear here have also been constructed independently in [45,46].

A Explicit computation of $c_b^{(r,s;r',s')}$

In this appendix we shall describe the strategy for explicit computation of the right hand side of (1.10)

$$\begin{aligned} F^{(r,s;r',s')}(\tau, z) &\equiv \frac{1}{MN} Tr_{RR;g_M^{r'}g_N^r}\left(g_M^{s'}g_N^s(-1)^{J_L+J_R}e^{2\pi i(\tau L_0-\bar\tau\bar L_0)}e^{2\pi i J_L z}\right) \\ &= \sum_{b=0}^{1}\sum_{j\in 2\mathbb{Z}+b, n\in\mathbb{Z}/MN} c_b^{(r,s;r',s')}(4n-j^2)e^{2\pi i(n\tau+jz)}\,, \end{aligned} \tag{A.1}$$

and hence of $c_b^{(r,s;r',s')}$. First of all $F^{(0,0;0,0)}$ is simply $1/MN$ times the elliptic genus of $\mathcal{M}$ and is given by [47]

$$\begin{aligned} F^{(0,0;0,0)}(\tau, z) &= 0 \quad \text{for } \mathcal{M} = T^4 \\ &= \frac{8}{MN}\left[\frac{\vartheta_2(\tau,z)^2}{\vartheta_2(\tau,0)^2}+\frac{\vartheta_3(\tau,z)^2}{\vartheta_3(\tau,0)^2}+\frac{\vartheta_4(\tau,z)^2}{\vartheta_4(\tau,0)^2}\right] \quad \text{for } \mathcal{M} = K3\,, \end{aligned} \tag{A.2}$$

where ϑ_i are the Jacobi theta functions. For non-vanishing r and/or r' the contribution to the right hand side of (A.1), coming from twisted sectors localized at the fixed points of $g_M^{r'}g_N^r$, can be computed by taking the size of K3 to be large so that the geometry near the fixed points is nearly flat. In this case the σ model near the fixed point can be regarded as the orbifold of a free field theory, and the action of $g_N^r g_M^{r'}$ near the fixed point may be represented by a rotation by some angle $2\pi\theta$ in one plane and rotation by an opposite angle $-2\pi\theta$ in the orthogonal plane. The action of $g_M^{s'} g_N^s$ near such a fixed point can be of two types, – either it takes the fixed point to a different fixed point or it leaves the fixed point fixed. In the former case the contribution to the trace in (A.1) vanishes, whereas in the latter case the action of $g_M^{s'} g_N^s$ near the fixed point can be represented by a rotation by $2\pi\phi$ in one plane and a rotation by $-2\pi\phi$ in the orthogonal plane. This gives a contribution to (A.1) of the form

$$\begin{aligned}
& \frac{1}{MN} \prod_{n=1}^{\infty} \Big\{ \left(1 - q^{n+\theta-1} e^{2\pi i\phi}\right)^{-2} \left(1 - q^{n-\theta} e^{-2\pi i\phi}\right)^{-2} \left(1 - q^{n+\theta-1} e^{2\pi i\phi} e^{2\pi i z}\right) \\
& \qquad \left(1 - q^{n+\theta-1} e^{2\pi i\phi} e^{-2\pi i z}\right) \left(1 - q^{n-\theta} e^{-2\pi i\phi} e^{2\pi i z}\right) \left(1 - q^{n-\theta} e^{-2\pi i\phi} e^{-2\pi i z}\right) \\
= & \frac{1}{MN} \frac{\vartheta_1(\tau, z+\theta\tau+\phi)\vartheta_1(\tau, -z+\theta\tau+\phi)}{\vartheta_1(\tau, \theta\tau+\phi)^2}\,, \qquad (A.3)
\end{aligned}$$

where $q \equiv e^{2\pi i\tau}$. The full contribution to $F^{(r,s;r',s')}(\tau, z)$ is obtained by summing over the contribution from all the fixed points of $g_N^r g_M^{r'}$ which are also fixed by $g_M^{s'} g_N^s$. Finally $F^{(0,s;0,s')}$ can be computed using the modular transformation rules [47]

$$F^{(r,s;r',s')}\left(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\right) = \exp\left[2\pi i \frac{cz^2}{c\tau+d}\right] F^{(cs+ar,ds+br;cs'+ar',ds'+br')}(\tau, z)\,. \qquad (A.4)$$

This gives a way to compute $F^{(r,s;r',s')}(\tau, z)$ using purely geometric data, namely the fixed points of the different elements of $\mathbb{Z}_M \times \mathbb{Z}_N$ and the action of the elements of $\mathbb{Z}_M \times \mathbb{Z}_N$ near the fixed points. Note that $F^{(r,s;r',s')}(\tau, z)$ constructed from (A.3), (A.4) is invariant under $z \to -z$. This is a consequence of the $SU(2)_L$ R-symmetry of the underlying conformal field theory that allows us to change the sign of J_L.

B Threshold Integral Representation of $\widetilde{\Phi}$

In this appendix we shall describe a threshold integral representation of $\widetilde{\Phi}$. For this we define

$$h_b^{(r,s;r',s')}(\tau) = \sum_{k \in \frac{1}{MN}\mathbb{Z} - \frac{b^2}{4}} c_b^{(r,s;r',s')}(4k) e^{2\pi i k\tau}, \qquad (B.1)$$

$$\Omega = \begin{pmatrix} \rho & v \\ v & \sigma \end{pmatrix}, \qquad (B.2)$$

and

$$\begin{aligned}
\frac{1}{2}p_R^2 &= \frac{1}{4\det Im\Omega} |-m_1\rho + m_2 + n_1\sigma + n_2(\sigma\rho - v^2) + jv|^2, \\
\frac{1}{2}p_L^2 &= \frac{1}{2}p_R^2 + m_1 n_1 + m_2 n_2 + \frac{1}{4}j^2\,. \qquad (B.3)
\end{aligned}$$

We now consider the 'threshold integral'

$$\widetilde{\mathcal{I}}(\rho,\sigma,v) = \sum_{r,s=0}^{N-1}\sum_{r',s'=0}^{M-1}\sum_{b=0}^{1}\widetilde{\mathcal{I}}_{r,s;r',s';b}\,, \tag{B.4}$$

where

$$\begin{aligned}\widetilde{\mathcal{I}}_{r,s;r',s';b} \;&=\; \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2}\Big[\sum_{\substack{m_1\in\mathbb{Z},m_2\in\mathbb{Z}/N,n_2\in N\mathbb{Z}-r\\ n_1\in\mathbb{Z}+\frac{r'}{M},j\in 2\mathbb{Z}+b}} q^{p_L^2/2}\bar{q}^{p_R^2/2}e^{2\pi i m_1 s'/M}e^{-2\pi i m_2 s}h_b^{(r,s;r',s')}(\tau) \\ &\quad -\delta_{b,0}\delta_{r,0}\delta_{r',0}c_0^{(0,s;0,s')}(0)\Big]\,,\end{aligned} \tag{B.5}$$

with

$$q \equiv e^{2\pi i\tau}\,. \tag{B.6}$$

$\mathcal{F}$ denotes the fundamental region of $SL(2,\ \mathbb{Z})$ in the upper half plane. The subtraction terms proportional to $c_0^{(0,s;0,s')}(0)$ have been chosen so that the integrand vanishes faster than $1/\tau_2$ in the $\tau\to i\infty$ limit, rendering the integral finite.

Following the manipulations outlined in [4] following earlier work of [48–50] one can show that

$$\widetilde{\mathcal{I}}(\rho,\sigma,v) = -2\ln[(\det Im\ \Omega)^{\widetilde{k}}] - 2\ln\widetilde{\Phi}(\rho,\sigma,v) - 2\ln\check{\Phi}(\bar\rho,\bar\sigma,\bar v) + \text{constant} \tag{B.7}$$

where $\widetilde{\Phi}$ has been defined in (1.9) and[6]

$$\begin{aligned}\check{\Phi}(\bar\rho,\bar\sigma,\bar v) \;&=\; e^{-2\pi i(\widetilde{\alpha}\bar\rho+\widetilde{\gamma}\bar\sigma+\widetilde{\beta}\bar v)}\prod_{b=0}^{1}\prod_{r=0}^{N-1}\prod_{r'=0}^{M-1}\prod_{\substack{k\in\mathbb{Z}+\frac{r'}{M},l\in\mathbb{Z},j\in 2\mathbb{Z}+b\\ k,l\geq 0,j<0\ for\ k=l=0}}\left[1-e^{-2\pi i r/N}\,e^{-2\pi i(k\bar\sigma+l\bar\rho+j\bar v)}\right]^a \\ a \;&\equiv\; \sum_{s=0}^{N-1}\sum_{s'=0}^{M-1}e^{-2\pi i(s'l/M+rs/N)}c_b^{(0,s;r',s')}(4kl-j^2)\,,\end{aligned} \tag{B.8}$$

$$\widetilde{k} \;=\; \frac{1}{2}\sum_{s=0}^{N-1}\sum_{s'=0}^{M-1}c_0^{(0,s;0,s')}(0)\,. \tag{B.9}$$

Since the detailed analysis of a specific integral of this type has been carried out in [4], we shall only describe the basic steps leading to (B.7), focussing on the main differences between the general case and the special case analyzed in [4].

---

[6] Since the exponent a in (1.9), (B.8) is a real number (in fact an integer) we have $\check{\Phi}(\bar\rho,\bar\sigma,\bar v) = \overline{\widetilde{\Phi}(\rho,\sigma,v)}$.

1. We first carry out Poisson resummation over m_1 and m_2 to express $\widetilde{\mathcal{I}}_{r,s;r',s';b}$ as

$$\int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2}\, \frac{Y}{\rho_2} \Big[N \sum_{\substack{k_1\in \mathbb{Z}+\frac{s'}{M},\, k_2\in N\mathbb{Z}-s \\ n_1\in\mathbb{Z}+\frac{r'}{M},\, n_2\in N\mathbb{Z}-r,\, j\in 2\mathbb{Z}+b}} h_b^{(r,s;r',s')}(\tau) e^{\mathcal{G}(\vec{n},\vec{k},j)} - \delta_{b,0}\delta_{r,0}\delta_{r',0} c_0^{(0,s;0,s')}(0) \Big] \tag{B.10}$$

where

$$\begin{aligned} \mathcal{G}(\vec{n},\vec{k},j) \;=\; & -\frac{\pi Y}{\rho_2^2\tau_2}|\mathcal{A}|^2 - 2\pi i\sigma \det A + \frac{\pi j}{\rho_2}(v\tilde{\mathcal{A}} - \bar{v}\mathcal{A}) - \frac{\pi n_2}{\rho_2}(v^2\tilde{\mathcal{A}} - \bar{v}^2\mathcal{A}) \\ & + \frac{2\pi i v_2^2}{\rho_2^2}(n_1 + n_2\bar{\rho})\mathcal{A} + 2\pi i\tau\frac{j^2}{4}\,, \end{aligned} \tag{B.11}$$

$$Y \equiv \rho_2\sigma_2 - v_2^2\,, \qquad A \equiv \begin{pmatrix} n_1 & k_1 \\ n_2 & k_2 \end{pmatrix}, \quad \mathcal{A} \equiv (1,\rho) A \begin{pmatrix} \tau \\ 1 \end{pmatrix}, \quad \tilde{\mathcal{A}} \equiv (1,\bar{\rho}) A \begin{pmatrix} \tau \\ 1 \end{pmatrix}. \tag{B.12}$$

2. Using (B.4), (B.10) and performing the sum over j using (A.1), (B.1) we get

$$\widetilde{\mathcal{I}} = \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \Big[N \sum_{k_1,n_1\in\frac{\mathbb{Z}}{M},\, k_2,n_2\in\mathbb{Z}} \mathcal{J}(A,\tau) - \sum_{s=0}^{N-1}\sum_{s'=0}^{M-1} c_0^{(0,s;0,s')}(0) \Big]\,, \tag{B.13}$$

where

$$\begin{aligned} \mathcal{J}(A,\tau) \;=\; & \frac{Y}{\rho_2} \exp\Big(-\frac{\pi Y}{\rho_2^2\tau_2}|\mathcal{A}|^2 - 2\pi i\sigma \det A \\ & - \frac{\pi n_2}{\rho_2}(v^2\tilde{\mathcal{A}} - \bar{v}^2\mathcal{A}) + \frac{2\pi i v_2^2}{\rho_2^2}(n_1 + n_2\bar{\rho})\mathcal{A} \Big) F^{(r,s;r',s')}\Big(\tau, -i\frac{v\tilde{\mathcal{A}} - \bar{v}\mathcal{A}}{2\,\rho_2}\Big) \\ & r = -n_2 \bmod N,\; s = -k_2 \bmod N,\; r' = M n_1 \bmod M,\; s' = M k_1 \bmod M\,. \end{aligned} \tag{B.14}$$

3. Using (A.4) one can show that

$$\mathcal{J}\left(A \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tau \right) = \mathcal{J}\left(A, \frac{a\tau+b}{c\tau+d} \right), \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\ \mathbb{Z})\,. \tag{B.15}$$

   Under the map $\tau \to (a\tau+b)/(c\tau+d)$ the fundamental region $\mathcal{F}$ of $SL(2,\ \mathbb{Z})$ gets mapped to its image. With the help of (B.15) we can restrict the sum over the matrix A to the following ranges [48]:

(a) Non-degenerate orbit:

$$A = \begin{pmatrix} k & m \\ 0 & p \end{pmatrix}, \quad k, m \in \frac{\mathbb{Z}}{M}, \quad p \in \mathbb{Z}, \quad p \neq 0, \quad 0 \leq m < k\,, \tag{B.16}$$

with the integration over τ ranging over two copies of the upper half plane.

(b) Degenerate orbit:

$$A = \begin{pmatrix} 0 & m \\ 0 & p \end{pmatrix}, \quad m \in \frac{\mathbb{Z}}{M}, \quad p \in \mathbb{Z}, \quad (m, p) \neq (0, 0)\,, \tag{B.17}$$

with the integration over τ ranging over the strip $-\frac{1}{2} \leq \tau_1 \leq \frac{1}{2}$, $\tau_2 > 0$.

(c) Zero orbit:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\,, \tag{B.18}$$

with the integration over τ ranging over the fundamental domain $\mathcal{F}$.

We shall briefly discuss the computation of the contribution from the non-degenerate orbits. The analysis of the contribution from the other orbits follows [4].

4. The steps needed for evaluating the contribution from the non-degenerate orbits are as follows:

   (a) We first expand $F^{(r,s;r',s')}(\tau, z)$ using (A.1) and change integration variable from τ_1 to τ_1':

   $$\tau_1' = \tau_1 + \frac{m}{k} + \frac{p}{k}\rho_1\,. \tag{B.19}$$

   Since for non-degenerate orbit the τ_1 integration ranges from $-\infty$ to ∞, τ_1' integral will also range from $-\infty$ to ∞.

   (b) After this change of variables the m dependence of the integral becomes simple. The explicit dependence on m takes the form $e^{-2\pi i m n/k}$, n and j being the integers appearing in the expansion (A.1). There is also an implicit dependence on m through $c_b^{(r,s;r',s')}(4n - j^2)$ since $s' = Mm \bmod M$ due to (B.14). The sum over m can be performed by expanding the summation range of m to $0 \leq m < Mk$ in steps of $1/M$, at the cost of having to divide the sum by a factor of M. We then sum over all values of m of the form $\frac{s'}{M}$ + integer for fixed s' and finally sum over integral s'

in the range $0 \leq s' \leq (M-1)$. For fixed s' the sum takes the form

$$\sum_{m\in \mathbb{Z}+\frac{s'}{M}, 0\leq m< Mk-1} e^{-2\pi i m n/k} = Mk\, e^{-2\pi i n s'/(Mk)} \sum_{l\in\mathbb{Z}} \delta_{n,kl} = Mk \sum_{l\in\mathbb{Z}} e^{-2\pi i l s'/M}\, \delta_{n,kl}\,. \tag{B.20}$$

The sum over n can now be performed using the Kronecker delta.

(c) Next we carry out the τ_1' integral which is Gaussian and the τ_2 integral using the identity

$$\int_0^\infty \frac{du}{u^{3/2}} e^{-au-bu^{-1}} = \sqrt{\frac{\pi}{b}} e^{-2\sqrt{ab}}\,. \tag{B.21}$$

At this stage the p dependent part of the summand takes the form

$$\frac{1}{|p|}\exp\left\{-2\pi i\sigma kp - 2\pi k|p|\sigma_2 - 2\pi kp\sigma_2 - 2\pi i lp\rho_1 - 2\pi l|p|\rho_2 - 2\pi i jpv_1 - 2\pi j|p|v_2\right\}\,. \tag{B.22}$$

(d) Finally we perform the sum over p by breaking it into contribution from $p>0$ and $p<0$ terms, making a change of variables $p\to -p$ in the $p<0$ terms, and using the identity

$$\begin{aligned}\sum_{p\in N\mathbb{Z}+s, p>0} \frac{1}{p} e^{2\pi i\alpha p} &= \frac{1}{N}\sum_{r=0}^{N-1}\sum_{p\in\mathbb{Z}, p>0} e^{2\pi i r(p-s)/N}\,\frac{1}{p}\, e^{2\pi i \alpha p} \\ &= -\frac{1}{N}\sum_{r=0}^{N-1} e^{-2\pi i r s/N}\, \ln(1-e^{2\pi i r/N} e^{2\pi i\alpha})\,. \end{aligned} \tag{B.23}$$

5. At the end of this manipulation we get the contribution from the non-degenerate orbits to be

$$\begin{aligned}-2\sum_{r,s=0}^{N-1}\sum_{r',s'=0}^{M-1}\sum_{b=0}^{1} \sum_{k\in\mathbb{Z}+\frac{r'}{M}, l\in\mathbb{Z}, j\in 2\mathbb{Z}+b, k>0, l\geq 0} & e^{-2\pi i rs/N} e^{-2\pi i l s'/M} c_b^{(0,s;r',s')}(4kl-j^2) \\ \left\{\ln\left[1-e^{2\pi i r/N} e^{2\pi i(k\sigma + l\rho + jv)}\right] + \ln\left[1-e^{-2\pi i r/N} e^{-2\pi i(k\bar\sigma + l\bar\rho + j\bar v)}\right]\right\} & + \text{constant}\,. \end{aligned} \tag{B.24}$$

One can recognize (B.24) as the contribution from $-2\ln\widetilde\Phi(\rho,\sigma,v) - 2\ln\check\Phi(\bar\rho,\bar\sigma,\bar v)$ except for the $k=0$ terms and the overall multiplicative factors in the product representation (1.9) of $\widetilde\Phi$ and (B.8) of $\check\Phi$. By carefully analyzing the contribution from the degenerate and the zero orbits one recovers the complete set of terms on the left hand side of (B.7).

C Zeroes of $\widetilde{\Phi}$

In this appendix we shall determine the locations of the zeroes of the function $\widetilde{\Phi}$. From (B.7) it follows that at the locations of the zeroes and poles of $\widetilde{\Phi}$ we have logarithmic divergences in $\widetilde{\mathcal{I}}$. Thus we can determine the locations of the zeroes and poles of $\widetilde{\Phi}$ by determining the singularities of $\widetilde{\mathcal{I}}$. Since the integrand in $\widetilde{\mathcal{I}}$ is finite for finite τ, and in its original form given in (B.4), (B.5) the integration over τ runs over the fundamental domain of $SL(2, \mathbb{Z})$, the only possible source of divergence is from the region of large τ_2. This requires that the powers of $q (\equiv e^{2\pi i\tau})$ and $\bar{q}$ be equal so that the τ_1 integral for large τ_2 does not vanish, and non-positive so that the τ_2 integral diverges. Since the only dependence on $\bar{q}$ is through the $\bar{q}^{p_R^2/2}$ term, and since p_R^2 is positive semi-definite, this requires p_R^2 to vanish. Furthermore p_L^2 is also positive semidefinite (although it is not directly apparent from its definition), and hence the only possible way to produce a divergence is to get a non-positive power of q from the expansion of $h_b^{(r,s;r',s')}(\tau)$.[7] Such terms are quite restricted, and one finds that the possible divergences in $\widetilde{\mathcal{I}}$ arise at [6, 40]

$$n_2(\rho\sigma - v^2) - m_1\rho + n_1\sigma + m_2 + jv = 0\,, \tag{C.1}$$

for values of (m_1, n_1, m_2, n_2, j) satisfying

$$\begin{aligned} & m_1 n_1 + m_2 n_2 + \frac{j^2}{4} = \frac{1}{4} - \beta, \qquad 0 \le \beta < \frac{1}{4}\,, \\ & m_1 \in \mathbb{Z}, \quad m_2 \in \mathbb{Z}/N, \quad n_2 \in \mathbb{Z}, \quad n_1 \in \mathbb{Z}/M, \quad j \in 2\,\mathbb{Z} + 1\,. \end{aligned} \tag{C.2}$$

At this point p_R^2 vanishes and $p_L^2/2 = (1/4) - \beta$ so that the $q^{p_L^2/2}\bar{q}^{p_R^2/2}$ factor becomes purely a function of τ and not of $\bar{\tau}$. This has to be cancelled against a similar τ dependent factor in h_b so that the τ_1 integral is finite and we get a logarithmically divergent τ_2 integral. The coefficient of the divergent term can be easily determined from (B.5), (B.1) and (B.7) and we get, near this point,[8]

$$\begin{aligned} \widetilde{\Phi}(\rho,\sigma,v) \;\sim\;& \left(n_2(\rho\sigma - v^2) - m_1\rho + n_1\sigma + m_2 + jv\right)^{\sum_{s'=0}^{M-1}\sum_{s=0}^{N-1} e^{2\pi i m_1 s'/M} e^{-2\pi i m_2 s} c_1^{(r,s;r',s')}(-1+4\beta)}\,, \\ & r = -n_2 \bmod N, \quad r' = M n_1 \bmod M\,. \end{aligned} \tag{C.3}$$

---

[7] There is a further restriction coming from the fact that for $p_L^2 = p_R^2 = 0$ all the m_i's, n_i's and j must vanish. The divergence in the τ_2 integral from such a term is removed by the subtraction term in (B.5). Thus a divergent τ_2 integral requires p_L^2 to be strictly positive. This explains the strict inequality $\beta < \frac{1}{4}$ in (C.2).

[8] We need to account for the fact that the terms in the expression for $\widetilde{\mathcal{I}}$ with (m_i, n_i, j) and $(-m_i, -n_i, -j)$ give identical results.

Let us first focus on the $r = r' = 0$ terms, i.e. configurations with $n_2 \in N\,\mathbb{Z}$, $n_1 \in \mathbb{Z}$. In this case it follows from (C.2) that β must vanish and so the right hand side of (C.3) involves $c_1^{(0,s;0,s')}(-1)$. Using (2.8) we now get, for $\mathcal{M} = K3$,

$$\sum_{s'=0}^{M-1}\sum_{s=0}^{N-1} e^{2\pi i m_1 s'/M} e^{-2\pi i m_2 s} c_1^{(0,s;0,s')}(-1) = \begin{cases} 2 \text{ for } m_1 \in M\,\mathbb{Z},\, m_2 \in \mathbb{Z} \\ 0 \text{ otherwise}\,, \end{cases} \tag{C.4}$$

and for $\mathcal{M} = T^4$,

$$\sum_{s'=0}^{M-1}\sum_{s=0}^{N-1} e^{2\pi i m_1 s'/M} e^{-2\pi i m_2 s} c_1^{(0,s;0,s')}(-1) = \begin{cases} 2 \text{ for } m_1 \in M\,\mathbb{Z},\, m_2 \in \mathbb{Z} \\ -1 \text{ for } m_1 \in M\,\mathbb{Z} \pm 1,\, m_2 \in \mathbb{Z} \mp \frac{1}{N} \\ 0 \text{ otherwise}\,. \end{cases} \tag{C.5}$$

In either case the exponent is positive, producing zeroes of $\widetilde{\Phi}$, only for $m_1 \in M\,\mathbb{Z}$ and $m_2 \in \mathbb{Z}$. Thus the net result is that the only zeroes of $\widetilde{\Phi}$ for $r = r' = 0$ are of the form:

$$\begin{aligned} \widetilde{\Phi}(\rho,\sigma,v) \quad &\sim \quad \left(n_2(\rho\sigma - v^2) - m_1\rho + n_1\sigma + m_2 + jv\right)^2\,, \\ &m_1 \in M\,\mathbb{Z}, \quad m_2 \in \mathbb{Z}, \quad n_1 \in \mathbb{Z}, \quad n_2 \in N\,\mathbb{Z}\,. \end{aligned} \tag{C.6}$$

Let us now analyze the contribution to the exponent in (C.3) when r and/or r' is non-zero. From the definition of the coefficients $c_b^{(r,s;r',s')}$ given in (1.10) it follows that the exponent in (C.3) can be interpreted as the number of states weighted by $(-1)^{J_L+J_R}$ in the sector twisted by $g_N^r g_M^{r'}$, and carrying g_M eigenvalue $e^{-2\pi i m_1/M}$, g_N eigenvalue $e^{2\pi i m_2}$, $J_L = \pm 1$, $L_0 = \beta$ and $\bar{L}_0 = 0$. If $g_M^{r'} g_N^r$ does not have a fixed point in $\mathcal{M}$ then this number is zero since by taking the size of $\mathcal{M}$ to be sufficiently large we can ensure that there will be no twisted sector state with $\bar{L}_0 = 0$. If $g_M^{r'} g_N^r$ has fixed points in $\mathcal{M}$, then we can compute this number by taking the size of $\mathcal{M}$ to be large so that near the fixed points we can regard the space as almost flat, with $g_N^r g_M^{r'}$ acting as rotation by some angle θ in one plane and by $-\theta$ in an orthogonal plane. Under such rotations all the bosons and fermions in the sigma model with target space $\mathcal{M}$ are twisted and hence there are no zero modes. Thus we have a unique ground state with ($J_L = J_R = 0$). Even after we apply left-moving oscillators to create states with $L_0 = \beta$, $J_L = \pm 1$, the states will continue to have $J_R = 0$. Thus the weight factor $(-1)^{J_L+J_R}$ is always -1 for $J_L = \pm 1$ and hence the exponent of (C.3) is always negative. This shows that (C.3) never gives a zero of $\widetilde{\Phi}$ for r and/or r' non-zero, and the only zeroes of $\widetilde{\Phi}$ are of the form given in (C.6).

References

[1] S. Chaudhuri, G. Hockney and J. D. Lykken, "Maximally Supersymmetric String Theories In D < 10," Phys. Rev. Lett. **75**, 2264 (1995) [arXiv:hep-th/9505054].

[2] S. Chaudhuri and J. Polchinski, "Moduli space of CHL strings," Phys. Rev. D **52**, 7168 (1995) [arXiv:hep-th/9506048].

[3] D. P. Jatkar and A. Sen, "Dyon spectrum in CHL models," JHEP **0604**, 018 (2006) [arXiv:hep-th/0510147].

[4] J. R. David, D. P. Jatkar and A. Sen, "Product representation of dyon partition function in CHL models," JHEP **0606**, 064 (2006) [arXiv:hep-th/0602254].

[5] A. Dabholkar and S. Nampuri, "Spectrum of dyons and black holes in CHL orbifolds using Borcherds lift," arXiv:hep-th/0603066.

[6] J. R. David and A. Sen, "CHL dyons and statistical entropy function from D1-D5 system," JHEP **0611**, 072 (2006) [arXiv:hep-th/0605210].

[7] J. R. David, D. P. Jatkar and A. Sen, "Dyon spectrum in N = 4 supersymmetric type II string theories," arXiv:hep-th/0607155.

[8] J. R. David, D. P. Jatkar and A. Sen, "Dyon spectrum in generic N = 4 supersymmetric Z(N) orbifolds," arXiv:hep-th/0609109.

[9] A. Dabholkar and D. Gaiotto, "Spectrum of CHL dyons from genus-two partition function," arXiv:hep-th/0612011.

[10] M. Cvetic and D. Youm, "Dyonic BPS saturated black holes of heterotic string on a six torus," Phys. Rev. D **53**, 584 (1996) [arXiv:hep-th/9507090].

[11] M. Cvetic and A. A. Tseytlin, "Solitonic strings and BPS saturated dyonic black holes," Phys. Rev. D **53**, 5619 (1996) [Erratum-ibid. D **55**, 3907 (1997)] [arXiv:hep-th/9512031].

[12] B. de Wit, "N = 2 electric-magnetic duality in a chiral background," Nucl. Phys. Proc. Suppl. **49**, 191 (1996) [arXiv:hep-th/9602060].

[13] B. de Wit, "N=2 symplectic reparametrizations in a chiral background," Fortsch. Phys. **44**, 529 (1996) [arXiv:hep-th/9603191].

[14] G. Lopes Cardoso, B. de Wit and T. Mohaupt, "Corrections to macroscopic supersymmetric black-hole entropy," Phys. Lett. B **451**, 309 (1999) [arXiv:hep-th/9812082].

[15] T. Mohaupt, “Black hole entropy, special geometry and strings,” Fortsch. Phys. **49**, 3 (2001) [arXiv:hep-th/0007195].

[16] R. Dijkgraaf, E. P. Verlinde and H. L. Verlinde, “Counting dyons in N = 4 string theory,” Nucl. Phys. B **484**, 543 (1997) [arXiv:hep-th/9607026].

[17] G. Lopes Cardoso, B. de Wit, J. Kappeli and T. Mohaupt, “Asymptotic degeneracy of dyonic N = 4 string states and black hole entropy,” JHEP **0412**, 075 (2004) [arXiv:hep-th/0412287].

[18] D. Shih, A. Strominger and X. Yin, “Recounting dyons in N = 4 string theory,” arXiv:hep-th/0505094.

[19] D. Gaiotto, “Re-recounting dyons in N = 4 string theory,” arXiv:hep-th/0506249.

[20] D. Shih and X. Yin, “Exact black hole degeneracies and the topological string,” JHEP **0604**, 034 (2006) [arXiv:hep-th/0508174].

[21] N. Banerjee, D. P. Jatkar and A. Sen, “Adding charges to N = 4 dyons,” arXiv:0705.1433 [hep-th].

[22] S. Banerjee, A. Sen and Y. K. Srivastava, “Generalities of Quarter BPS Dyon Partition Function and Dyons of Torsion Two,” arXiv:0802.0544 [hep-th].

[23] S. Banerjee, A. Sen and Y. K. Srivastava, “Partition Functions of Torsion > 1 Dyons in Heterotic String Theory on T^6,” arXiv:0802.1556 [hep-th].

[24] A. Dabholkar, J. Gomes and S. Murthy, “Counting all dyons in N =4 string theory,” arXiv:0803.2692 [hep-th].

[25] A. Sen, “Quantum Entropy Function from AdS(2)/CFT(1) Correspondence,” Int. J. Mod. Phys. A **24**, 4225 (2009) [arXiv:0809.3304 [hep-th]].

[26] A. Sen, “Arithmetic of Quantum Entropy Function,” JHEP **0908**, 068 (2009) [arXiv:0903.1477 [hep-th]].

[27] A. Sen, “A Twist in the Dyon Partition Function,” arXiv:0911.1563 [hep-th].

[28] M. G. Alford, J. March-Russell and F. Wilczek, "DISCRETE QUANTUM HAIR ON BLACK HOLES AND THE NONABELIAN AHARONOV-BOHM EFFECT," Nucl. Phys. B **337**, 695 (1990).

[29] S. R. Coleman, J. Preskill and F. Wilczek, "Quantum hair on black holes," Nucl. Phys. B **378**, 175 (1992) [arXiv:hep-th/9201059].

[30] A. Gregori, E. Kiritsis, C. Kounnas, N. A. Obers, P. M. Petropoulos and B. Pioline, "R**2 corrections and non-perturbative dualities of N = 4 string ground states," Nucl. Phys. B **510**, 423 (1998) [arXiv:hep-th/9708062].

[31] E. Kiritsis, "Introduction to non-perturbative string theory," arXiv:hep-th/9708130.

[32] N. Banerjee, D. P. Jatkar and A. Sen, "Asymptotic Expansion of the N=4 Dyon Degeneracy," JHEP **0905**, 121 (2009) [arXiv:0810.3472 [hep-th]].

[33] S. Murthy and B. Pioline, "A Farey tale for N=4 dyons," JHEP **0909**, 022 (2009) [arXiv:0904.4253 [hep-th]].

[34] S. Chaudhuri and D. A. Lowe, "Type IIA heterotic duals with maximal supersymmetry," Nucl. Phys. B **459**, 113 (1996) [arXiv:hep-th/9508144].

[35] P. S. Aspinwall, "Some relationships between dualities in string theory," Nucl. Phys. Proc. Suppl. **46**, 30 (1996) [arXiv:hep-th/9508154].

[36] A. Sen, "Walls of marginal stability and dyon spectrum in N = 4 supersymmetric string theories," arXiv:hep-th/0702141.

[37] A. Dabholkar, D. Gaiotto and S. Nampuri, "Comments on the spectrum of CHL dyons," arXiv:hep-th/0702150.

[38] A. Sen, "Two Centered Black Holes and N=4 Dyon Spectrum," arXiv:0705.3874 [hep-th].

[39] M. C. N. Cheng and E. Verlinde, "Dying Dyons Don't Count," arXiv:0706.2363 [hep-th].

[40] A. Sen, "Black Hole Entropy Function, Attractors and Precision Counting of Microstates," arXiv:0708.1270 [hep-th].

[41] C. N. Pope, "Axial Vector Anomalies And The Index Theorem In Charged Schwarzschild And Taub - Nut Spaces," Nucl. Phys. B **141**, 432 (1978).

[42] J. P. Gauntlett, N. Kim, J. Park and P. Yi, “Monopole dynamics and BPS dyons in N = 2 super-Yang-Mills theories,” Phys. Rev. D **61**, 125012 (2000) [arXiv:hep-th/9912082].

[43] D. Gaiotto, A. Strominger and X. Yin, “New connections between 4D and 5D black holes,” JHEP **0602**, 024 (2006) [arXiv:hep-th/0503217].

[44] R. Dijkgraaf, G. W. Moore, E. P. Verlinde and H. L. Verlinde, “Elliptic genera of symmetric products and second quantized strings,” Commun. Math. Phys. **185**, 197 (1997) [arXiv:hep-th/9608096].

[45] S. Govindarajan, talk at NSM 2010, IIT Bombay.

[46] S. Govindarajan, ”BKM Lie superalgebras from twisted CHL dyons”, to appear.

[47] T. Kawai, Y. Yamada and S. K. Yang, “Elliptic Genera And N=2 Superconformal Field Theory,” Nucl. Phys. B **414**, 191 (1994) [arXiv:hep-th/9306096].

[48] L. J. Dixon, V. Kaplunovsky and J. Louis, “Moduli dependence of string loop corrections to gauge coupling constants,” Nucl. Phys. B **355**, 649 (1991).

[49] T. Kawai, “$N = 2$ heterotic string threshold correction, $K3$ surface and generalized Kac-Moody superalgebra,” Phys. Lett. B **372**, 59 (1996) [arXiv:hep-th/9512046].

[50] C. D. D. Neumann, “The elliptic genus of Calabi-Yau 3-folds and 4-folds, product formulae and generalized Kac-Moody algebras,” J. Geom. Phys. **29**, 5 (1999) [arXiv:hep-th/9607029].

arXiv:1208.6254v3 [math.RT] 20 Aug 2016

GENERALIZED MOONSHINE IV: MONSTROUS LIE ALGEBRAS

SCOTT HUAI-LEI CARNAHAN

ABSTRACT. For each element of the Fischer-Griess Monster sporadic simple group, we construct an infinite dimensional Lie algebra equipped with a projective action of the centralizer of that element. Our construction is given by a string-theoretic "add a spacetime torus and quantize" functor applied to an abelian intertwining algebra that is formed from a family of twisted modules of the Monster vertex operator algebra. We prove that for all Fricke elements in the Monster, the characters of centralizers acting on the corresponding irreducible twisted modules are Hauptmoduln. From these results, we resolve Norton's Generalized Moonshine Conjecture.

CONTENTS

1. Introduction 1
1.1. Generalized Moonshine 1
1.2. Overview 6
1.3. Acknowledgments 7
2. Vertex algebraic input 7
2.1. Vertex algebras and twisted modules 7
2.2. Abelian intertwining algebras 8
2.3. Unrolling 12
2.4. Adding a spacetime torus 16
2.5. Projective action of centralizer 17
3. Lie algebras 19
3.1. Virasoro preliminaries 19
3.2. Old covariant quantization 22
3.3. BRST cohomology 24
3.4. Borcherds-Kac-Moody property 27
4. Comparison theorems 30
4.1. Lie algebra comparison 30
4.2. Main theorem 31
References 32

1. INTRODUCTION

For each element g of the Monster, we construct an infinite dimensional Lie algebra equipped with a canonical projective action of the centralizer of g. For those g satisfying the Fricke property, we show that the Lie algebras are Borcherds-Kac-Moody with a real simple root, and we use the resulting twisted denominator identities to prove Norton's Generalized Moonshine Conjecture for the class of irreducible twisted $V^\natural$-modules.

1.1. Generalized Moonshine.

We give a brief reminder of monstrous moonshine before describing its generalization. In the late 1970s, McKay observed that the q^1-term in the q-expansion of the normalized modular function $J = q^{-1} + 196884q + 21493760q^2 + \cdots$ is one more than the dimension of the smallest nontrivial complex representation of the Fischer-Griess Monster simple group $\mathbb{M}$, which is a sporadic simple group of order about 8×10^{53}. This observation suggested a relationship between the J-function and representations of the Monster that was further reinforced by additional computations by McKay and Thompson. Conway and Norton then did a massive computation with characters, producing evidence that there is an infinite dimensional graded representation $V = \bigoplus_{n=0}^{\infty} V_n$ of the Monster, such that the graded trace $T_g(\tau) = \sum_{n=0}^{\infty} \mathrm{Tr}(g|V_n)q^{n-1}$ of any element is the Fourier expansion (under $q = e^{2\pi i\tau}$) of a holomorphic function on the complex upper half-plane $\mathfrak{H}$ that is moreover a Hauptmodul [Conway-Norton-1979]. That is, for each $g \in \mathbb{M}$, there is a group $\Gamma_g \subset SL_2(\mathbb{R})$ under which $T_g(\tau)$ is invariant, such that the quotient curve $X_g = \mathfrak{H}/\Gamma_g$ is genus zero, and T_g generates the field of meromorphic functions on X_g. A candidate representation $V^\natural$ was constructed by Frenkel, Lepowsky, and Meurman, who showed that it is a vertex operator algebra whose automorphism group is precisely $\mathbb{M}$, and its graded dimension T_1 is J [Frenkel, et al, 1988]. Borcherds then proved Conway and Norton's conjecture for $V^\natural$ [Borcherds-1992].

In the Conway-Norton paper, the authors suggest that there may be moonshine-like phenomena for groups other than $\mathbb{M}$, and shortly afterward, Queen did some computations that yielded concrete evidence that one can obtain Hauptmoduln from characters of a variety of large finite groups [Queen-1981]. In particular, the low-order coefficients of many such modular functions can be assembled from representations of sporadic simple groups other than the Monster.

For illumination, we consider a level 2 example. The irreducible representations of the Baby Monster sporadic simple group $\mathbb{B}$ have dimension $1, 4371, 96255, \ldots$, while the power series $q^{-1} + 4372q + 96256q^2 + \cdots$ is the q-expansion of $\frac{\Delta(\tau)}{\Delta(2\tau)} + \frac{\Delta(2\tau)}{\Delta(\tau)} + 24$, which is a Hauptmodul for $\Gamma_0(2)^+$. It is difficult to resist the temptation to decompose the coefficients as small combinations of dimensions of irreducible representations: $4372 = 4371 + 1$ and $96256 = 96255 + 1$. Furthermore, we encounter promising evidence in the form of the character of an element in conjugacy class 2A: the first term yields $q^{-1} - 492q + O(q^2)$, and there is a well-known modular function $\sqrt{j(2\tau) - 1728} = q^{-1} - 492q - 22590q^3 - \cdots$ that is a Hauptmodul for $\Gamma_0(2|2)$. However, a complication arises: the q^2 term in our character does not vanish, since the trace of 2A on $\underline{96255}$ is 4863, not -1. Despite the initial promising evidence, we find that the Baby Monster group does not yield moonshine-like behavior. However, there is a central extension $2.\mathbb{B}$, with an irreducible representation of dimension 96256, where the trace of 2A is zero. By replacing the Baby Monster with its central extension, we eliminate this obstruction to assembling the characters into modular functions. Similar behavior appears when considering the sporadic group Fi_{24}, where one needs to pass to a triple cover to get a representation of dimension 783, yielding a level 3 modular function $q^{-1} + 783q + \cdots$. What these examples have in common is that both $2.\mathbb{B}$ and $3.Fi_{24}$ are centralizers of elements of the Monster.

Norton employed these computations and many of his own to formulate the following conjecture:

Conjecture. *(Generalized Moonshine [Norton-1987], revised in [Norton-2001]) There exists a rule that assigns to each element g of the Monster simple group $\mathbb{M}$ a graded projective representation $V(g) = \bigoplus_{n \in \mathbb{Q}} V(g)_n$ of the centralizer $C_{\mathbb{M}}(g)$, and to each pair (g, h) of commuting elements of $\mathbb{M}$ a holomorphic function $Z(g, h, \tau)$ on the complex upper half-plane $\mathfrak{H}$, satisfying the following conditions:*

(1) *There is some lift $\tilde{h}$ of h to a linear transformation on $V(g)$ such that*

$$Z(g, h, \tau) = \sum_{n \in \mathbb{Q}} Tr(\tilde{h}|V(g)_n)q^{n-1}.$$

(2) *$Z(g, h, \tau)$ is invariant up to constant multiplication under simultaneous conjugation of the pair (g, h) in $\mathbb{M}$.*

(3) *$Z(g, h, \tau)$ is either constant, or a Hauptmodul for some genus zero congruence group.*

(4) *For any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, $Z(g^a h^c, g^b h^d, \tau)$ is proportional to $Z(g, h, \frac{a\tau + b}{c\tau + d})$.*

(5) $Z(g,h,\tau) = J(\tau) = \frac{E_4(\tau)^3}{\Delta(\tau)} - 744 = q^{-1} + 196884q + \cdots$ *if and only if* $g = h = 1 \in \mathbb{M}$.

This is a generalization of the Monstrous Moonshine Conjecture, in the sense that fixing $g = 1$ yields an assertion that one has a graded representation of $\mathbb{M}$ whose characters form Hauptmoduln. It also subsumes the example observation above, in the sense that if $q^{-1}+4372q+96256q^2+\cdots$ is the Moonshine character of an element g in class 2A of the Monster, then $Z(g,1,\tau)$ is proportional to $q^{-1/2}+4372q^{1/2}+96256q+\cdots$, and these coefficients decompose into well-behaved representations of $2.\mathbb{B}$, the centralizer of an element in class 2A of the Monster.

In contrast to the situation with the original Monstrous Moonshine conjecture, we do not have a list of conjugacy classes of commuting pairs of elements in $\mathbb{M}$, or a list of character tables of central extensions of centralizers of elements at our disposal. Conversations with group theorists suggest that these explicit data are still beyond current technology, so at the moment, we cannot solve the conjecture by computational brute force. Instead, one needs to find an interpretation of the spaces $V(g)$ and the functions Z, and one must seek general methods for proving that traces like Z yield Hauptmoduln, and for showing that traces are $SL_2(\mathbb{Z})$-compatible.

Shortly after Norton published this conjecture, Dixon, Ginsparg, and Harvey found a physical interpretation for claims 1,2,4, and 5: For each $g \in \mathbb{M}$, the space $V(g)$ is the "Hilbert space twisted by g" in a conformal field theory with $\mathbb{M}$ symmetry, and $Z(g,h)$ is a genus one partition function for a torus twisted in "space" by g and in "time" by h [Dixon-Ginsparg-Harvey-1988]. In algebraic language, we would say that for each $g \in \mathbb{M}$, we have a distinguished irreducible g-twisted module $V^\natural(g)$ of the Monster vertex operator algebra, and one may hope that such objects exist and are unique up to isomorphism.

The first major mathematical breakthrough toward a solution of this conjecture then came in [Dong-Li-Mason-2000], with the following results:

1. Theorem 10.3 asserts that for any $g \in \mathbb{M}$, there exists an irreducible g-twisted $V^\natural$-module, and it is unique up to isomorphism. By Schur's Lemma, such a module $V^\natural(g)$ admits a canonical projective action of $C_{\mathbb{M}}(g)$. By general nonsense about conjugation, this resolves Claim 2, at least on the level of formal power series.
2. Theorem 8.1 implies the series $\sum_{n\in\mathbb{Q}} Tr(\tilde{h}|V(g)_n)q^{n-1}$ converges absolutely uniformly on compact subsets of the complex upper half-plane to a holomorphic function, for any lift $\tilde{h}$ of an element $h \in C_{\mathbb{M}}(g)$. We can therefore define $Z(g,h,\tau)$ to be this function for some choice of $\tilde{h}$. This resolves Claim 1, and translates the power-series version of Claim 2 to a holomorphic function version.
3. Theorem 10.3 also contains an assertion that implies $Z(g,1,\tau)$ is a constant multiple of $Z(1,g,-1/\tau)$. This resolves Claim 5, essentially by the fact that $Z(1,g,\tau)$ has level greater than 1 for any $g \neq 1$.
4. Theorem 10.1 reduces Claim 4 to an conjecture about g-rationality of $V^\natural$, i.e., that any g-twisted $V^\natural$-module is isomorphic to a direct sum of copies of $V^\natural(g)$.

Claim 4 was recently resolved: By Corollary 5.26 of [Carnahan-Miyamoto-2016], g-rationality holds for any simple regular non-negatively graded vertex operator algebra with an invariant bilinear form, and $V^\natural$ satisfies these properties. The final remaining claim is Claim 3, the Hauptmodul-or-constant property, and this is the subject of this paper.

Under the current interpretation, Claims 1,2,4, and 5 are in fact general claims about non-negatively graded holomorphic C_2-cofinite vertex operator algebras and their automorphisms. That is, these claims hold true for an infinite class of examples, and don't have much to do with the Monster, except for the fact that $V^\natural$ is a particularly well-behaved vertex operator algebra. Claim 3 has a more exceptional, moonshine-like quality, since there are only finitely many Hauptmoduln, and the restriction on level gives an indication of smallness. For this reason, a solution cannot be expected to appear directly from general facts about vertex operator algebras or conformal field theory. Perhaps unsurprisingly, this is the only claim that was not given a physical interpretation by Dixon-Ginsparg-Harvey. Proving such an exceptional claim seems to require a specialized strategy, and this is where we come to our second major breakthrough.

In an unpublished preprint [Höhn-2003b], Höhn resolved the Hauptmodul claim for the case where g is an element in the conjugacy class 2A. The resolution of this case is by itself a substantial advance, but more importantly, Höhn's preprint contained a general strategy for proving that Generalized Moonshine trace functions are Hauptmoduln. Alternatively, one may say that Höhn worked out precisely the ingredients needed to extend Borcherds's proof of the Monstrous Moonshine Conjecture to the general case in question. We shall call this the Borcherds-Höhn program:

1. Build an abelian intertwining algebra structure on the direct sum $\bigoplus_{i=0}^{|g|-1} V^\natural(g^i)$.
2. Add a spacetime torus: Make a conformal vertex algebra of central charge 26 by taking a graded tensor product with the abelian intertwining algebra of a suitable hyperbolic lattice.
3. Quantize: Apply a bosonic string quantization functor to the conformal vertex algebra. This yields a Lie algebra $\mathfrak{m}_g$ equipped with a canonical projective action of $C_{\mathbb{M}}(g)$ by automorphisms.
4. Generate a Lie algebra L_g whose denominator identity is some automorphic infinite product. This Lie algebra has a "nice shape", in the sense that its simple roots and homology are well-controlled.
5. Comparison: Show that the Lie algebra $\mathfrak{m}_g$ is isomorphic to L_g. By *transport de structure*, we obtain a Lie algebra with both a group action and nice shape.
6. Hauptmodul conclusion: Use the twisted denominator identity to produce recursion relations on the characters that are strong enough to conclude that the characters are Hauptmoduln.

When $g = 1$, this program gives an outline of Borcherds's proof [Borcherds-1992], with the first step given by Frenkel-Lepowsky-Meurman's construction of $V^\natural$. When g is an element of type 2A, this program gives an outline of Höhn's proof of Generalized Moonshine for $2.\mathbb{B}$ [Höhn-2003b]. In both of those cases, the last step was settled by combining recursion relations with an explicit computation of the first few terms, to match the proposed functions with the actual characters. In general, we do not have character tables for central extensions of centralizers in the Monster, so such explicit computations are out of the question. Instead, we found a general Hauptmodul criterion involving equivariant Hecke operators in [Carnahan-2010], generalizing the results of [Cummins-Gannon-1997]. While Cummins-Gannon could be applied to remove the computations from Borcherds's proof of the Monstrous Moonshine conjecture [Borcherds-1992], our methods could be used to eliminate explicit computations for all cases we expected to encounter, including Höhn's proof. This settles the last step, conditional on the constructions in the previous steps.

The step with L_g was resolved in [Carnahan-2012], i.e., we constructed automorphic infinite products and candidate Lie algebras for each element of the Monster. We also defined a precise condition, called "Fricke-compatibility", on a projective action of $C_{\mathbb{M}}(g)$ on L_g that suffices for showing that the Hauptmodul condition holds for g of Fricke type. The question of producing such an action is essentially equivalent to the question of constructing $\mathfrak{m}_g$ and showing that $\mathfrak{m}_g \cong L_g$, and this is precisely what we resolve in this paper.

The first step is the most technically challenging. For example, in the case $g = 1$, this is the construction of the moonshine module $V^\natural$, which takes up the whole book [Frenkel, et al, 1988]. Before my recent project with Miyamoto culminating in [Carnahan-Miyamoto-2016], I had formed an elaborate plan to construct suitable abelian intertwining algebras using the theory of conformal blocks on twisted nodal curves together with some work eliminating homological obstructions. Some of this theory was sketched in my dissertation [Carnahan-2007], and the third paper in this series

[Carnahan≥2016] is intended to present a full development. I believe the methods are still useful in a general mathematical sense, and the first installment of results should appear in the near future.

We therefore come to the third major breakthrough, which rendered my plan superfluous for the purpose of proving Generalized Moonshine. In [van Ekeren, et al, 2015], the authors showed that there is an abelian intertwining algebra structure on a sum of twisted modules of a well-behaved vertex operator algebra, by an ingenious application of the Verlinde formula [Huang-2008] together with the fixed-point regularity results of [Carnahan-Miyamoto-2016]. Furthermore, they determined the precise characters of modules of the fixed point vertex algebra $(V^\natural)^{\langle g \rangle}$. Thus, their results essentially establish the first step, and go a long way toward the comparison step.

We now describe how the results of this paper fit in. In order to construct $\mathfrak{m}_g$, we need a functorial way to pass from abelian intertwining algebras to Lie algebras. We will employ an "add a spacetime torus and quantize" functor that has existed in some form for over 40 years, but we supply some necessary details that have not appeared in the literature. Adding a torus is just a degree-wise tensor product with the abelian intertwining algebra of a well-chosen hyperbolic rational lattice. The quantization functor takes conformal vertex algebras of central charge 26 to Lie algebras. We exploit the "cancellation of oscillators" phenomenon to describe the homogeneous spaces of the Lie algebras explicitly in terms of the input object. This theory was one of the main engines behind Frenkel's bound on root multiplicities of the Lie algebra whose simple roots form the Leech Lattice as its Dynkin diagram [Frenkel-1985], Borcherds's proof of the Monstrous Moonshine conjecture, and Hoehn's proof of Generalized Moonshine in the 2A case.

Once $\mathfrak{m}_g$ is constructed, we show that it is a Borcherds-Kac-Moody algebra by introducing a new criterion that is particularly well-suited to considering the output of the quantization functor. Then, the homogeneous subspaces become root spaces, and they are canonically identified with certain subspaces of twisted modules as representations of some central extension of $C_{\mathbb{M}}(g)$. The comparison isomorphism $\mathfrak{m}_g \cong L_g$ then follows from a comparison of root multiplicities, and this is made possible by the precise character work in [van Ekeren, et al, 2015].

The reader may be interested to know just how much information about $V^\natural$ and the Monster was necessary for our proof to work. In fact, we need almost no fine structure at all. By examining the proof in this paper, together with the work on which it depends, we see that in fact we have the following result:

Theorem 1.1.1. *Let V be a holomorphic C_2-cofinite vertex operator algebra of central charge 24 whose character is $J(\tau)$, and let G be a finite order group of automorphisms of V. Then any rule assigning to each $g \in G$ the irreducible g-twisted module $V(g)$, and to each commuting pair (g,h) the function $Z(g,h;\tau) = \mathrm{Tr}(\tilde{h} q^{L_0 - 1} | V(g))$ for some lift $\tilde{h} \in \widetilde{C_G(g)}$ of h to a linear transformation on $V(g)$, satisfies the conditions of Norton's conjecture, with G in place of $\mathbb{M}$.*

This is quite similar to the situation with Borcherds's proof of the Monstrous moonshine conjectures - one only needs central charge 24, an invariant bilinear form, and the character J to get completely replicable functions (which were later shown to be either highly degenerate or Hauptmoduln in [Cummins-Gannon-1997]). However, without the construction of $V^\natural$ in [Frenkel, et al, 1988], this result loses most of its value, since the connection to the Monster is lost. As it happens, Frenkel-Lepowsky-Meurman conjectured in section 3 of the introduction of *loc. cit.* that any V satisfying (more or less) the hypotheses in this theorem is isomorphic to $V^\natural$, so if we assume a suitable variant of the FLM Uniqueness Conjecture, then the statement given here becomes essentially the same as our Main Theorem (Theorem 4.2.2).

We note that there is previous work on Generalized Moonshine that establishes some Hautpmodul results for twisted modules - perhaps the most advanced was [Ivanov-Tuite-2002]. However, these results were conditional on both the Uniqueness Conjecture and the construction of holomorphic orbifolds established in [van Ekeren, et al, 2015].

We do not consider this the end of the Generalized Moonshine story. There are conjectural refinements of our main theorem, and intriguing conjectural relations between Generalized Moonshine and other fields of mathematics and physics. We focus on three further questions:

1. Norton has suggested stronger versions of this conjecture based on computational evidence [Norton-2001]. In particular, he asserts that the constant ambiguity can be reduced to a root of unity (in particular a 24th root), and that for any (g, h), there is some rational k such that each coefficient of $Z(g, h, \tau + k)$ generates a field extension of $\mathbb{Q}$ with Galois group of exponent 2. Terry Gannon has proposed a further refinement of the root-of-unity restriction, asserting that the proportionality constants are completely controlled by a 3-cocycle in the cohomology of $\mathbb{M}$, with coefficients in $U(1)$. The precise conjectural relationship between a suitable 3-cocycle and the proportionality constants is detailed in section 2 of [Gaberdiel-Persson-Volpato-2013]. Such a cocycle is expected from orbifold conformal field theory considerations [Dijkgraaf-Pasquier-Roche-1990], and an analogue of this refinement for M_{24} moonshine was established in [Gaberdiel, et al, 2012]. We note that this refinement is a general statement about twistings of holomorphic C_2-cofinite vertex operator algebras, and does not have a substantial connection to the Hauptmodul problem we solve in this paper.
2. Recall that almost all of the claims in the Generalized Moonshine conjecture were given physical interpretations in [Dixon-Ginsparg-Harvey-1988], and the remaining claim concerns Hauptmoduln. As it happens, there is a conjectural connection between Hauptmoduln and physics through Rademacher sums, as suggested in [Duncan-Frenkel-2011]. The general picture is that AdS/CFT correspondence seems to relate 2+1-dimensional gravity in the extreme quantum regime with conformal field theories satisfying smallness conditions. In particular, Duncan and Frenkel propose the existence of a class of second-quantized twisted chiral quantum gravity theories that can be manipulated to produce the monstrous Lie algebras $\mathfrak{m}_g$ constructed in this paper.
3. There are intriguing conjectural connections between moonshine and equivariant elliptic cohomology. The folklore idea is that elliptic cohomology is somehow controlled by, or is a shadow of, conformal field theory. Concretely, an element of the elliptic cohomology of the classifying space of a finite group G is a datum that includes a family of modular forms parametrized by a pair of commuting elements of G, satisfying an $SL_2(\mathbb{Z})$-compatibility condition. It was noted quite early (and I am not sure who was first - perhaps Mason) that the weight zero version of this condition is precisely given as claim 4 in the Generalized Moonshine conjecture. While elliptic cohomology does not play a direct role in our proof of the Generalized Moonshine conjecture, it was essential in an inspirational sense. Indeed, equivariant Hecke operators, seen as power operations in equivariant elliptic cohomology, were the key to the formulation of a good Hauptmodul criterion in [Carnahan-2010]. I believe the precise formula was first written in [Ganter-2009], where one may also find a topological interpretation of replicability. This piece of history, together with hints relating the formation of $\mathfrak{m}_g$ to exponential operations, suggests that the Hauptmodul claim may be amenable to a topological interpretation.

We would be interested in seeing any developments toward the refined versions of this conjecture, and the establishment of stronger connections to other mathematical disciplines.

1.2. **Overview.** In section 2, we introduce some vertex algebraic constructions, such as abelian intertwining algebras and the "add a torus" functors. The main purpose of this section is to take the results of [van Ekeren, et al, 2015], and transmute their construction into a form that is amenable to quantization.

In section 3, we construct Lie algebras, using a generalization of Borcherds's method for constructing the Monster Lie algebra. Due to the functorial nature of the construction, we obtain canonical group actions by homogeneous automorphisms. We show that for Fricke elements, the resulting Lie algebras are necessarily Borcherds-Kac-Moody, and we show that the root multiplicities match the dimensions of certain eigenspaces in twisted modules.

In section 4, we compare root multiplicities. For Fricke elements, this then yields isomorphisms with the Lie algebras constructed in [Carnahan-2012], and we find that characters of centralizers are Hauptmoduln. This then implies the full conjecture.

1.3. **Acknowledgments.** The author would like to thank Richard Borcherds, Jethro van Ekeren, Jeff Harvey, Yoshitake Hashimoto, Gerald Höhn, Satoshi Kondo, Gregory Moore, Luboš Motl, Yu Nakayama, Simon Norton, and Nils Scheithauer for helpful conversations and advice. Kazuya Kawasetsu, Sven Möller, and Hiromichi Yamada offered suggestions and corrections to an earlier version of this article.

This research was partly supported by NSF grant DMS-0354321, JSPS Kakenhi Grant-in-Aid for Young Scientists (B) 24740005, the Program to Disseminate Tenure Tracking System, MEXT, Japan, and the World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan.

2. Vertex algebraic input

Algebraic structures related to vertex algebras are essential to moonshine, but we will not need to spend much time on fine details, since most of the technical aspects have been worked out. We will review some basic facts, then pursue some consequences of the work of [van Ekeren, et al, 2015].

We will use the notation $e(s)$ to denote the normalized exponential $e^{2\pi i s}$.

2.1. Vertex algebras and twisted modules.

Definition 2.1.1. A vertex algebra over $\mathcal{C}$ is a complex vector space V equipped with a distinguished vector $\mathbf{1} \in V$, a distinguished linear transformation $T : V \to V$, and a left multiplication map $Y : V \to (\operatorname{End} V)[[z, z^{-1}]]$, written $Y(a,z) = \sum_{n \in \mathbb{Z}} a_n z^{-n-1}$ satisfying the following conditions:

1. $Y(\mathbf{1}, z) = id_V = id_V z^0$, and $Y(a,z)\mathbf{1} \in a + zV[[z]]$.
2. For any $a, b \in V$, $Y(a,z)b \in V((z))$, i.e., $a_n b = 0$ for n sufficiently large.
3. $[T, Y(a,z)] = \frac{d}{dz} Y(a,z)$
4. The Jacobi identity: for any $a, b \in V$,

$$x^{-1}\delta\left(\frac{y-z}{x}\right) Y(a,y)Y(b,z) - x^{-1}\delta\left(\frac{z-y}{-x}\right) Y(b,z)Y(a,y) = z^{-1}\delta\left(\frac{y-x}{z}\right) Y(Y(a,x)b,z)$$

where $\delta(z) = \sum_{n \in \mathbb{Z}} z^n$ and $\delta(\frac{y-z}{x})$ is expanded as a formal power series in z.

An automorphism of a vertex algebra V is a linear transformation $g : V \to V$ that fixes $\mathbf{1}$, and satisfies $Y(ga,z)gb = gY(a,z)b$ for all $a, b \in V$.

Definition 2.1.2. A conformal vertex algebra of central charge $c \in \mathbb{C}$ is a vertex algebra equipped with a distinguished nonzero vector ω, satisfying the following conditions:

1. If we write $Y(\omega, z) = \sum_{n \in \mathbb{Z}} L_n z^{-n-2}$, then the coefficients $L_n \in \operatorname{End} V$ satisfy the Virasoro relations:
$$[L_m, L_n] = (m-n)L_{m+n} + c\frac{m^3 - m}{12}\delta_{m+n,0} id_V$$
2. L_0 acts semisimply on V, with integer eigenvalues.
3. $L_{-1} = T$.

An automorphism of a conformal vertex algebra is an automorphism of the underlying vertex algebra that fixes ω. A vertex operator algebra is a conformal vertex algebra for which the eigenvalues of L_0 are bounded below and have finite multiplicity.

We will only consider one vertex operator algebra in this paper, namely the Monster vertex operator algebra $V^\natural$ that was constructed in [Frenkel, et al, 1988].

Vertex algebras have a more complicated representation theory than ordinary algebras, because modules may be twisted by automorphisms of the vertex algebra.

Definition 2.1.3. Let V be a vertex algebra over $\mathbb{C}$, and let g be an automorphism of order n. Then a g-twisted V-module is a complex vector space M equipped with an action map $Y^M : V \to (\operatorname{End} M)[[z^{1/n}, z^{-1/n}]]$, written $Y^M(a,z) = \sum_{k \in \frac{1}{n}\mathbb{Z}} a_k z^{-k-1}$ satisfying the following conditions:

1. $Y^M(\mathbf{1}, z) = id_M$
2. For any $a \in V$, $u \in M$, $Y^M(a,z)u \in M((z^{1/n}))$, i.e., $a_k u = 0$ if k is sufficiently large.
3. If $a \in V$ satisfies $ga = e(k/n)a$ for some $k \in \mathbb{Z}$, then $Y^M(a,z) \in z^{k/n}(\operatorname{End} M)[[z, z^{-1}]]$.
4. The twisted Jacobi identity holds: if $ga = e(k/n)a$, then for any $b \in V$,

$$x^{-1}\delta\left(\frac{y-z}{x}\right)Y^M(a,y)Y^M(b,z) - x^{-1}\delta\left(\frac{z-y}{-x}\right)Y^M(b,z)Y^M(a,y) =$$
$$= z^{-1}\left(\frac{y-x}{z}\right)^{-k/n}\delta\left(\frac{y-x}{z}\right)Y^M(Y(a,x)b,z)$$

If V is a vertex operator algebra, these conditions define the notion of "weak g-twisted V-module". An "admissible g-twisted V-module" is a weak g-twisted V-module that admits a $\frac{1}{n}\mathbb{Z}_{\geq 0}$-grading that is compatible with the L_0-grading on V. An "ordinary g-twisted V-module" is an admissible g-twisted V-module for which L_0 acts semisimply, with finite dimensional eigenspaces, and eigenvalues that are bounded below in each coset of $\mathbb{Z}$ in $\mathbb{C}$. When $g = 1$, we replace "g-twisted V-module" with "V-module".

Definition 2.1.4. Here are a few more technical conditions about vertex algebras. We won't use them in any specific way, but they appear as conditions in the theorems we need.

1. A vertex algebra V is C_2-cofinite if the subspace $C_2(V)$ spanned by $\{a_2 b | a, b \in V\}$ satisfies the property that $V/C_2(V)$ is finite dimensional.
2. A vertex operator algebra V is g-rational if all admissible g-twisted V-modules are direct sums of irreducible ordinary g-twisted V-modules.
3. A vertex operator algebra V is holomorphic if all admissible V-modules are isomorphic to direct sums of V.
4. A vertex operator algebra V is of CFT type if L_0 has only non-negative eigenvalues, and the kernel of L_0 is spanned by $\mathbf{1}$.

Remark 2.1.5. Theorem 10.3 in [Dong-Li-Mason-2000] asserts that if a vertex operator algebra V is C_2-cofinite and holomorphic, then there exists a unique irreducible g-twisted V-module, which we will write as $V(g)$. For any automorphism h of V, one may change the action of V on the underlying vector space of $V(g)$ to produce a module isomorphic to $V(h^{-1}gh)$. By Schur's Lemma, we obtain a canonical projective action of $C_{\operatorname{Aut} V}(g)$ on the underlying vector space of $V(g)$, which is compatible with the action on V (but it is not an action by twisted module automorphisms, since those are just scalars).

Definition 2.1.6. We write $\operatorname{Aut} V(g)$ to denote the group of vector space automorphisms of $V(g)$ that are compatible with automorphisms of V - it is a central extension of $C_{\operatorname{Aut} V}(g)$ by $\mathbb{C}^\times$.

2.2. Abelian intertwining algebras.

Dong and Lepowsky introduced abelian intertwining algebras essentially as vertex algebras with some monodromy obstructing commutativity. We describe them briefly, and we describe the theorem we need from [van Ekeren, et al, 2015].

Definition 2.2.1. ([MacLane-1952]) Let A be an abelian group. The Eilenberg-Mac Lane cochain complex with coefficients in $\mathbb{C}^\times$ has the following description in low degree:

$$K^1 \xrightarrow{d^1} K^2 = \{f : A \times A \to \mathbb{C}^\times\} \xrightarrow{d^2} K^3 = \{(F,\Omega) : A^{\oplus 3} \times A^{\oplus 2} \to \mathbb{C}^\times\} \xrightarrow{d^3} \cdots$$

where:

1. $K^1 = \{\phi : A \to \mathbb{C}^\times\}$ and $K^2 = \{f : A \times A \to \mathbb{C}^\times\}$ are groups of set-theoretic maps.
2. $K^3 = \{(F,\Omega) : A^{\oplus 3} \times A^{\oplus 2} \to \mathbb{C}^\times\}$ is a group of pairs of set-theoretic maps.
3. The map $d^1 : K^1 \to K^2$ is given by the usual group cohomology coboundary $\phi \mapsto d^1\phi$, defined by $d^1\phi(i,j) = \phi(j) - \phi(i+j) + \phi(i)$.
4. The map $d^2 : K^2 \to K^3$ is given by the group cohomology coboundary $d^2 f(i,j,k) = f(j,k) - f(i+j,k) + f(i,j+k) - f(i,j)$ together with the antisymmetrizer: $(i,j) \mapsto \frac{f(i,j)}{f(j,i)}$.
5. The map $d^3 : K^3 \to K^4$ vanishes if and only if the following conditions are satisfied:
   (a) $F(i,j,k)F(i,j+k,\ell)F(j,k,\ell) = F(i+j,k,\ell)F(i,j,k+\ell)$ for all $i,j,k,\ell \in A$
   (b) $F(i,j,k)^{-1}\Omega(i,j+k)F(j,k,i)^{-1} = \Omega(i,j)F(j,i,k)^{-1}\Omega(i,k)$
   (c) $F(i,j,k)\Omega(i+j,k)F(k,i,j) = \Omega(j,k)F(i,k,j)\Omega(i,k)$

The kernel of d^n is written $Z^n_{ab}(A,\mathbb{C}^\times)$, and elements are called abelian n-cocycles. The cohomology of the Eilenberg-MacLane complex is called abelian cohomology, and written $H^n_{ab}(A,\mathbb{C}^\times)$. An abelian 3-cocycle is normalized if $F(i,j,0) = F(i,0,k) = F(0,j,k) = 1$ and $\Omega(i,0) = \Omega(0,j) = 1$ for all $i,j,k \in A$.

We note that any cohomology class has a normalized representative.

Lemma 2.2.2. *([MacLane-1952] Theorem 3) If (F,Ω) is an abelian 3-cocycle, then the map $Q : A \to \mathbb{C}^\times$ defined by $i \mapsto \Omega(i,i)$ is a quadratic function, i.e., $Q(i) = Q(-i)$ for all $i \in A$, and $\frac{Q(i+j+k)Q(i)Q(j)Q(k)}{Q(i+j)Q(i+k)Q(j+k)} = 1$ for all $i,j,k \in A$. Furthermore, the trace map $(F,\Omega) \mapsto (i \mapsto \Omega(i,i))$ induces a bijection from $H^3_{ab}(A,\mathbb{C}^\times)$ to the set of $\mathbb{C}^\times$-valued quadratic functions on A.*

Proof. This is proved as [Eilenberg-Mac Lane-1954], theorem 26.1. □

Definition 2.2.3. Let A be an abelian group, let (F,Ω) be an abelian 3-cocycle on A. We write $q_\Omega : A \to \mathbb{C}/\mathbb{Z}$ for the unique quadratic map such that $e(q_\Omega(a)) = \Omega(a,a)$ for all $a \in A$, and we write $b_\Omega : A \times A \to \mathbb{C}/\mathbb{Z}$ for the induced bilinear form.

We take the following definition from [Dong-Lepowsky-1993]

Definition 2.2.4. Let A be an abelian group, let (F,Ω) be a normalized abelian 3-cocycle on A. Then an abelian intertwining algebra of level $N \in \mathbb{Z}_{\geq 1}$ and central charge $c \in \mathbb{Q}$ associated to the datum (A,F,Ω) is a complex vector space V equipped with

1. a $\frac{1}{N}\mathbb{Z} \times A$-grading $V = \bigoplus_{n \in \frac{1}{N}\mathbb{Z}} V_n = \bigoplus_{i \in A} V^i = \bigoplus_{n \in \frac{1}{N}\mathbb{Z}, i \in A} V^i_n$
2. a left-multiplication $Y : V \to (\operatorname{End} V)[[z^{1/N}, z^{-1/N}]]$ written $Y(a,z) = \sum_{n \in \frac{1}{N}\mathbb{Z}} a_n z^{-n-1}$, and
3. distinguished vectors $\mathbf{1} \in V^0_0$ and $\omega \in V^0_2$

satisfying the following conditions for any $i,j,k \in A$, $a \in V^i$, $b \in V^j$, $u \in V^k$, and $n \in \frac{1}{N}\mathbb{Z}$:

1. $a_n b \in V^{i+j}$.
2. $a_n b = 0$ for n sufficiently large.
3. $Y(\mathbf{1},z)a = a$
4. $Y(a,z)\mathbf{1} \in a + zV[[z]]$.
5. $Y(a,z)b = \sum_{k \in b_\Omega(i,j)+\mathbb{Z}} a_k b z^{-k-1}$.

(6) The Jacobi identity:

$$x^{-1}\left(\frac{y-z}{x}\right)^{b_\Omega(i,j)}\delta\left(\frac{y-z}{x}\right)Y(a,y)Y(b,z)u$$
$$-B(i,j,k)x^{-1}\left(\frac{z-y}{e^{i\pi}x}\right)^{b_\Omega(i,j)}\delta\left(\frac{z-y}{-x}\right)Y(b,z)Y(a,y)u=$$
$$=F(i,j,k)z^{-1}\delta\left(\frac{y-x}{z}\right)Y(Y(a,x)b,z)\left(\frac{y-x}{z}\right)^{-b_\Omega(i,k)}u$$

where $B(i,j,k)=\frac{\Omega(i,j)F(i,j,k)}{F(j,i,k)}$.

(7) The coefficients of $Y(\omega,z)=\sum_{k\in\mathbb{Z}}L_kz^{-k-2}$ satisfy the Virasoro relations at central charge c.

(8) $L_0a=na$ if $a\in V_n$.

(9) $\frac{d}{dz}Y(a,z)=Y(L_{-1}a,z)$.

Given a suitably nondegenerate abelian intertwining algebra (namely one where left multiplication by elements of V^i is not the zero map), one has an additional $\mathbb{C}/\mathbb{Z}$-valued quadratic form q_Δ on A, given by setting $q_\Delta(i)$ to be the conformal weight of V^i. It is not hard to show that the associated bilinear forms satisfy $b_\Omega=-b_\Delta$ (see, e.g., the discussion in [van Ekeren, et al, 2015] after Theorem 4.2). Furthermore, it is often the case that we have $q_\Omega=-q_\Delta$ - when this holds, we say that the abelian intertwining algebra is "even".

Dong and Lepowsky give the following generalization of the lattice vertex algebra construction: for any finite rank free abelian group L with a $\frac{1}{N}\mathbb{Z}$-valued quadratic form, we may produce the abelian intertwining algebra $V=V_L$ attached to L. It is L-graded, and the L-homogeneous pieces are the usual modules $\pi_\lambda^{L\otimes\mathbb{C}}$ of the Heisenberg vertex algebra $\pi_0^{L\otimes\mathbb{C}}$ for the quadratic vector space $L\otimes\mathbb{C}$ attached to vectors $\lambda\in L$. This abelian intertwining algebra is even, i.e., it satisfies $q_\Delta=-q_\Omega$. One needs to choose a cocycle (F,Ω) in the cohomology class of q_Ω to obtain an abelian intertwining algebra, but we have the following way to pass between different choices.

Lemma 2.2.5. *Given an abelian intertwining algebra V with cocycle (F,Ω), and any function $f\in K^2$ satisfying $f(i,0)=f(0,j)=1$, the tuple $(V,1,\omega,f\cdot Y)$ is an abelian intertwining algebra with cocycle $d_2f\cdot(F,\Omega)$. Here, $f\cdot Y$ is the map $V\to(\operatorname{End}V)[[z^{1/N},z^{-1/N}]]$ that takes a pair $a\in V^i$, $b\in V^j$ to $f(i,j)\cdot Y(a,z)b$.*

Proof. This is Remark 12.23 in [Dong-Lepowsky-1993]. ☐

Definition 2.2.6. Let q be a quadratic form on an abelian group A that takes values in $\frac{1}{N}\mathbb{Z}/\mathbb{Z}$. We define $AIA^c_{(A,q)}$ to be the following groupoid: objects are abelian intertwining algebras of level N with central charge c, associated to the datum (A,F,Ω) for some abelian 3-cocycle (F,Ω) such that $q_\Omega=e(q)$. Morphisms are homogeneous isomorphisms of abelian intertwining algebras.

By Lemma 2.2.5, this groupoid is equipped with an action of the abelian 3-coboundary group by autofunctors.

The last structure we introduce is a graded tensor product.

Definition 2.2.7. Let V be an object in $AIA^c_{(A,q)}$ with cocycle (F,Ω), and let W be an object in $AIA^{c'}_{(A,q')}$ with cocycle (F',Ω'). Then we define $V\otimes^A W$ to be the following object in $AIA^{c+c'}_{(A,qq')}$ with cocycle $(FF',\Omega\Omega')$:

(1) The underlying graded vector space is $\bigoplus_{i\in A}V^i\otimes W^i$

(2) Multiplication is induced by tensor product maps. That is, $Y(a\otimes a',z)(b\otimes b')=Y(a,z)b\otimes Y(a',z)b'$ for $a,b\in V$, and $a',b'\in W$.

(3) $\mathbf{1}_{V\otimes^A W}=\mathbf{1}_V\otimes\mathbf{1}_W$

(4) $\omega_{V\otimes^A W} = \omega_V \otimes \mathbf{1}_W + \mathbf{1}_V \otimes \omega_W$

The following theorem describes the results of van Ekeren, Möller, and Scheithauer that we need.

Theorem 2.2.8. *Let V be a holomorphic C_2-cofinite vertex operator algebra of CFT type, let g be an automorphism of V of order $n \in \mathbb{Z}_{\geq 1}$, and let $G = \langle g \rangle$ be the cyclic group generated by g. Suppose the L_0-spectrum of the irreducible twisted module $V(g^i)$ is strictly positive whenever $g^i \neq 1$. Then there is a distinguished element $t \in \mathbb{Z}/n\mathbb{Z}$ and a unique assignment of homomorphisms $\phi_i : G \to \operatorname{Aut} V(g^i)$ satisfying the following conditions:*

1. *If we let $V^{(i,j)} = \{v \in V(g^i) | \phi_i(g)v = e(j/n)v\}$, then the set $\{V^{(i,j)}\}_{i,j\in\mathbb{Z}/n\mathbb{Z}}$ forms a complete list of isomorphism classes of irreducible V^G-modules. They are simple currents, whose fusion rules satisfy $V^{(i,j)} \boxtimes V^{(k,l)} \cong V^{(i+k,j+l+c_{2t}(i,k))}$, where c_{2t} is the "add with carry" 2-cocycle $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by $c_{2t}(i,k) = \begin{cases} 0 & i_n + k_n < n \\ 2t & i_n + k_n \geq n \end{cases}$. Here, i_n and k_n denote the unique elements of $\{0, \dots, n-1\}$ to which i and k reduce modulo n, respectively, and we let D denote the central extension of $\mathbb{Z}/n\mathbb{Z}$ by $\mathbb{Z}/n\mathbb{Z}$ defined by the 2-cocycle c_{2t}.*
2. *There exists an abelian intertwining algebra ${}^gV \cong \bigoplus_{i\in\mathbb{Z}/n\mathbb{Z}} V(g^i) \cong \bigoplus_{i,j\in\mathbb{Z}/n\mathbb{Z}} V^{(i,j)}$, whose multiplication is compatible with the fusion rules on $V^{(i,j)}$, graded by the abelian group D, with $q_\Omega(i,j) = -\frac{ij}{n} - \frac{i_n^2 t_n}{n^2}$. Furthermore, ${}^gV$ is even, i.e., $q_\Delta(i,j) = -q_\Omega(i,j)$.*

Furthermore, if we write

$$T(v,i,j,\tau) = \operatorname{Tr}(o(v)\phi_i(g^j)q^{L_0-c/24}|V(g^i))$$
$$T_{(i,j)}(v;\tau) = \operatorname{Tr}(o(v)q^{L_0-c/24}|V^{(i,j)})$$

for all $v \in V^G$, $i,j \in \mathbb{Z}/n\mathbb{Z}$ and $q = e(\tau)$, where $o(v)$ is the zero-mode map (which is $v_{wt(v)-1}$ for homogeneous v), we obtain holomorphic functions on $\mathfrak{H}$ that satisfy the following modular transformation rules:

$$T(v,i,j,\tau+1) = e(\frac{i_n^2 t_n}{n^2} - \frac{c}{24})T(v,i,j+i,\tau)$$
$$T(v,i,j,\frac{-1}{\tau}) = \tau^{wt[v]}e(\frac{-2ti_nj_n}{n^2})T(v,j,-i,\tau)$$
$$T_{(i,j)}(v;\tau+1) = e(\frac{ij}{n} + \frac{i_n^2 t_n}{n^2} - \frac{c}{24})T_{(i,j)}(v;\tau)$$
$$T_{(i,j)}(v;\frac{-1}{\tau}) = \frac{1}{n}\tau^{wt[v]} \sum_{k,l\in\mathbb{Z}/n\mathbb{Z}} e(-\frac{kj+il}{n})e(\frac{-2ti_nk_n}{n^2})T_{(k,l)}(v;\tau)$$

where $wt[v]$ denotes the conformal weight with respect to Zhu's bracket grading [Zhu-1996].

Proof. We point out where the specific claims are given in [van Ekeren, et al, 2015]. The identification of eigenspaces with irreducible V^G-modules is Proposition 5.2, and the fact that they are simple currents is Proposition 5.6.

The fusion rules, conformal weights and S-matrix entries are given in Theorem 5.12. The modular transformation rules for $T(v,i,j,\tau)$ and $T_{(i,j)}(v;\tau)$ then follow from the identification of S-matrix entries and the conformal weights of modules.

The existence and properties of the abelian intertwining algebra are given in Theorem 4.1. □

We now show that this theorem can be applied to the object of our interest. Recall that an element g in the Monster is called Fricke if its McKay-Thompson series $T_g(\tau)$ is proportional to $T_g(-1/N\tau)$ for some N. Note that this is a definition that depends on the construction of the moonshine module $V^\natural$ in [Frenkel, et al, 1988]. By Borcherds's proof of the Monstrous Moonshine

Conjecture [Borcherds-1992], we have the following necessary and sufficient conditions for g to be Fricke: T_g has a pole at zero, or $T_g(\tau)$ is equal to $T_g(-1/N\tau)$ for some N. The Fricke property is conjugation-invariant, since the McKay-Thompson series is a trace. The conjectural enumeration of functions in [Conway-Norton-1979] (proved to be valid by Borcherds) shows that 120 of the 171 distinct McKay-Thompson series are Fricke, and 143 of the 194 conjugacy classes in the Monster are Fricke.

Theorem 2.2.9. *The Monster vertex operator algebra $V^\natural$ is holomorphic, C_2-cofinite, and of CFT type. For any nontrivial element g of the Monster, the L_0-spectrum of $V(g)$ is strictly positive. That is, the conditions of Theorem 2.2.8 are satisfied for any automorphism of $V^\natural$.*

Proof. The holomorphic property is proved in [Dong-1994], the C_2-cofinite property is proved in [Dong-Li-Mason-2000], and CFT type is clear from the construction [Frenkel, et al, 1988]. By Theorem 13.1 of [Dong-Li-Mason-2000], the character of $V^\natural(g)$ is a constant multiple of $T_g(-1/\tau)$, where $T_g(\tau) = Tr(gq^{L_0-1}|V^\natural)$ is the McKay-Thompson series. If T_g is regular at $\tau = 0$ (i.e., g is non-Fricke), then the lowest-weight space of $V(g)$ has L_0-eigenvalue at least 1. If T_g has a pole at $\tau = 0$ (i.e., g is Fricke), then the pole has order $1/N$, where N is the level of T_g, and the lowest-weight space of $V^\natural(g)$ has L_0-eigenvalue $1 - 1/N$. We conclude that as long as g is nontrivial, the lowest weight space has positive L_0-eigenvalue. ☐

We may then determine the precise graded dimension of twisted modules of $V^\natural$ - this eliminates the scalar ambiguity in [Dong-Li-Mason-2000].

Corollary 2.2.10. *For any $g \in \mathbb{M}$, hypothesis A_g from section 4.4 of [Carnahan-2012] is satisfied, i.e., $\sum_{n\in\mathbb{Q}} \dim V^\natural(g)_n q^{n-1} = T_g(-1/\tau)$.*

Proof. The left side of the equation is $T(\mathbf{1}, 1, 0, \tau)$, and the right side is $T(\mathbf{1}, 0, 1, -1/\tau)$, as defined in Theorem 2.2.8. By Theorem 2.2.9, we may apply Theorem 2.2.8 to $V^\natural$ and its twisted modules, so the two quantities are equal. ☐

Corollary 2.2.11. *If $V = V^\natural$, and $g \in \mathbb{M}$ has eigengroup $\Gamma_0(n|h) + e_1, e_2, \dots$ (in the notation of [Conway-Norton-1979]), then the distinguished element $t \in \mathbb{Z}/n\mathbb{Z}$ given in Theorem 2.2.8 is equal to $-n/h$ if g is Fricke and n/h if g is non-Fricke.*

Proof. The value of t is uniquely determined by the transformation $\tau \mapsto \tau + n$ on the graded dimension of $V^\natural(g)$. Theorem 2.2.8 gives us $T(\mathbf{1}, 1, 0, \tau - n) = e(-\frac{t_n}{n})T(\mathbf{1}, 1, 0, \tau)$, and Section 5 of [Conway-Norton-1979] gives us the transformation rule $T_g(\tau)|_0 \left(\begin{smallmatrix} 1 & 0 \\ n & 1 \end{smallmatrix}\right) = e(\pm 1/h)T_g(\tau)$, where we have a plus sign if and only if g is Fricke. The latter Möbius transformation conjugates to the $\tau \mapsto \tau - n$ transformation on $T(\mathbf{1}, 1, 0, \tau)$, so we have $t_n = \mp n/h$. ☐

2.3. **Unrolling.** In Section 3.1 of [Carnahan-2012] we introduce vector-valued modular functions $\hat{F}$ and F, which are uniquely determined by the $SL_2(\mathbb{Z})$ action on McKay-Thompson series. In section 4.1, we will need to match the functions $T(\mathbf{1}, i, j, \tau)$ with $\hat{F}_{i,j}(\tau)$ and $T_{(i,j)}(\tau)$ with $F_{i,j}(\tau)$ in order to show that certain Lie algebras are isomorphic. However, this is only possible when $t = 0$ given our current setup, because that is the only way the modular transformation rules can match. We will fix this problem by "unrolling" the abelian intertwining algebra into a $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$-graded object in a unique way, for N any multiple of $\frac{n^2}{(n,t)}$. By passing to the unrolled object, the modular transformation rules match in all cases.

Lemma 2.3.1. *Let $V_A = \bigoplus_{a\in A} V_a$ be an abelian intertwining algebra with cocycle (F_A, Ω_A). Let (B, q) be an abelian group with $\mathbb{C}/\mathbb{Z}$-valued quadratic form, and let $\pi : (B, q) \to (A, (a \mapsto \Omega_A(a, a)))$ be the quotient by a null subgroup $I \subset B^\perp$. Then there is a unique "pullback" abelian intertwining algebra $W_B = \bigoplus_{b\in B} W_b$ satisfying the following properties:*

(1) $W_b = \begin{cases} V_a & b \in B, \pi(b) = a \\ 0 & b \notin B \end{cases}$
(2) *The multiplication map* $W_b \otimes W_{b'} \to W_{b+b'}\{z\}$ *is given by the multiplication* $V_a \otimes V_{a'} \to V_{a+a'}\{z\}$ *from* V_A *when* $b, b' \in B$, $\pi(b) = a$, *and* $\pi(b') = a'$.

The cocycle (F_B, Ω_B) *attached to* W_B *is precisely the pullback of the cocycle* (F_A, Ω_A) *attached to* V_A.

Proof. The fact that W_B satisfies the axioms of an abelian intertwining algebra follows straightforwardly from our definition of the multiplication operation - the Jacobi identity for W_B reduces immediately to the Jacobi identity for V_A. ☐

Lemma 2.3.2. *If* $i : (A, q_A) \to (B, q_B)$ *is a quadratic embedding (i.e.,* i *is an injective homomorphism of abelian groups, and* $q_B \circ i = q_A$*), and* V_A *is an* (A, q_A)*-graded abelian intertwining algebra, then there is a unique "extension by zero" abelian intertwining algebra* $W_B = \bigoplus_{b\in B} W_b$ *satisfying the following properties*

(1) $W_b = \begin{cases} V_a & b = i(a) \\ 0 & b \notin i(A) \end{cases}$
(2) *The multiplication map* $W_b \otimes W_{b'} \to W_{b+b'}\{z\}$ *is given by the multiplication* $V_a \otimes V_{a'} \to V_{a+a'}\{z\}$ *from* V_A *when* $b = i(a)$, *and* $b' = i(a')$. *For all other values of* b *and* b', *at least one of the input spaces is zero, so multiplication is trivial.*

The cocycle (F_B, Ω_B) *attached to* W_B *is uniquely determined by* (F_A, Ω_A) *up to a coboundary that vanishes identically on* $i(A)$.

Proof. Omitted. ☐

Definition 2.3.3. Given a diagram $(A, q) \xleftarrow{\pi} (A', q') \xhookrightarrow{i} (A'', q'')$ of quadratic spaces, we define the unrolling functor $AIA^c_{(A,q)} \to AIA^c_{(A'',q'')}$ as the composite of the pullback given in Lemma 2.3.1 and the extension by zero given in Lemma 2.3.2.

The following proposition shows that we can make the modular transformation rules in Theorem 2.2.8 somewhat cleaner by unrolling.

Proposition 2.3.4. *Let* V *be a holomorphic* C_2*-cofinite vertex operator algebra of CFT type, let* g *be an automorphism of* V *of order* $n \in \mathbb{Z}_{\geq 1}$, *such that the* L_0*-spectrum of the irreducible twisted module* $V(g^i)$ *is strictly positive whenever* $g^i \neq 1$. *Let* $t \in \mathbb{Z}/n\mathbb{Z}$ *be the distinguished element given by Theorem 2.2.8, and let* $N = \frac{rn^2}{(n,t)}$ *for some* $r \in \mathbb{Z}_{\geq 1}$. *Finally, let* B *denote the group* $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ *with quadratic form* $(a, b) \mapsto e(\frac{ab}{N})$. *Then the subgroup* $B' \subset B$ *generated by* $(0, \frac{rn}{(n,t)})$ *and* $(1, \frac{t_n r}{(n,t)})$ *admits a quadratic-form-preserving quotient map* $B' \to D$ *by* $a(1, \frac{t_n r}{(n,t)}) + b(0, \frac{rn}{(n,t)}) \mapsto (a_n, b_n + 2t_n\lfloor a/n \rfloor)$, *and there exists a unique set of homomorphisms* $\{\psi_a : \mathbb{Z}/N\mathbb{Z} \to \operatorname{Aut} V(g^a)\}_{a=0}^{N-1}$ *such that the abelian intertwining algebra* ${}^g_N W \cong \bigoplus_{(a,b)\in B} W^{(a,b)} \cong \bigoplus_{a=0}^{N-1} V(g^a)$ *unrolled from* ${}^gV$ *satisfies the following conditions:*

(1) *For each* $a, b \in \mathbb{Z}/N\mathbb{Z}$, *we have* $W^{(a,b)} = \{v \in V(g^a) | \psi_a(1)v = e(\frac{b}{N})v\}$, *and when* $W^{(a,b)} \neq 0$, *it is a simple current (in particular, an irreducible module) for the fixed-point algebra* V^g.
(2) *For any* $a, b \in \mathbb{Z}/N\mathbb{Z}$, *we have the equality* $\psi_a(b) = \phi_a(g^b)e(b\frac{a_n t_n}{n^2} - b\frac{t_n\lfloor a/n \rfloor}{n})$ *of automorphisms of* $V(g^a)$.
(3) *If we define* $T(a, b, \tau) = \operatorname{Tr}(\psi_a(b)q^{L_0 - c/24}|V(g^a))$ *and* $T_{(a,b)}(\tau) = \operatorname{Tr}(q^{L_0-c/24}|W^{(a,b)})$ *for all* $a, b \in \mathbb{Z}/N\mathbb{Z}$ *and* $q = e(\tau)$, *then we obtain holomorphic functions on* $\mathfrak{H}$ *that satisfy the*

following modular transformation rules:

$$\begin{aligned} T(a,b,\tau+1) &= e(-\frac{c}{24})T(a,b+a,\tau) \\ T(a,b,\frac{-1}{\tau}) &= T(b,-a,\tau) \\ T_{(a,b)}(\tau+1) &= e(\frac{ab}{N}-\frac{c}{24})T_{(a,b)}(\tau) \\ T_{(a,b)}(\frac{-1}{\tau}) &= \frac{1}{N}\sum_{p,q\in\mathbb{Z}/n\mathbb{Z}} e(-\frac{bp+aq}{N})T_{(p,q)}(\tau) \end{aligned}$$

In particular (following [Carnahan-2012] Lemma 3.6), if c is a multiple of 24, then the collection $\{T_{(a,b)}\}_{a,b\in\mathbb{Z}/N\mathbb{Z}}$ forms a vector-valued modular function of type ρ_M for $M = II_{1,1}(N)$, and the collection $\{T(a,b,\tau)\}_{a,b\in\mathbb{Z}/N\mathbb{Z}}$ describes a function on the moduli space $\mathcal{M}_{Ell}^{\mathbb{Z}/N\mathbb{Z}}$ of elliptic curves with $\mathbb{Z}/N\mathbb{Z}$-torsors.

Proof. We first check that $B' \to D$ preserves the quadratic form: $q_{B'}(a(1,\frac{t_n r}{(n,t)}) + b(0,\frac{rn}{(n,t)})) = \frac{ab}{n} + \frac{a^2 t_n}{n^2} + \mathbb{Z}$, while $q_D(a_n, b_n + 2t_n\lfloor a/n\rfloor) = \frac{a_n b_n + 2a_n t_n \lfloor a/n\rfloor}{n} + \frac{a_n^2 t_n}{n^2} + \mathbb{Z}$. Because $a = n\lfloor a/n\rfloor + a_n$, we have $\frac{a^2 t_n}{n^2} + \mathbb{Z} = \frac{a_n^2 t_n}{n^2} + \frac{2a_n t_n\lfloor a/n\rfloor}{n} + \mathbb{Z}$, so the two expressions agree.

We therefore have a well-defined unrolling functor, and an abelian intertwining algebra ${}^g_N W$ whose B-homogeneous components and multiplication maps are either pulled back from ${}^g V$ or zero. The homomorphisms ψ_a are uniquely determined by the B-grading on ${}^g_N W$. The simple current property follows from the first listed claim in 2.2.8.

We now check that $\psi_a(b) = \phi_a(g^b)e(b\frac{a_n t_n}{n^2} - b\frac{t_n\lfloor a/n\rfloor}{n})$ on $V(g^a)$. For any $a(1,\frac{t_n r}{(n,t)}) + j(0,\frac{rn}{(n,t)}) \in B'$, we have $\psi_a(1)|_{W^{a(1,\frac{t_n r}{(n,t)})+j(0,\frac{rn}{(n,t)})}} = e(\frac{at_n}{n^2} + \frac{j}{n})$. The module is the pullback of $V^{(a_n, j_n + 2t_n\lfloor a/n\rfloor)}$, where $\phi_a(g)$ acts as $e(\frac{j_n + 2t_n\lfloor a/n\rfloor}{n})$. Dividing the two answers, and taking the bth power yields the answer we want.

It remains to consider the modular transformation properties. We reduce the first two equations to the last two equations by noting that $\{T(a,b,\tau)\}$ and $T_{(a,b)}(\tau)$ are related by a discrete Fourier transform. That is, the equations

$$T(a,b,\tau) = \sum_{k\in\mathbb{Z}/N\mathbb{Z}} e\left(\frac{bk}{N}\right) T_{(a,k)}(\tau)$$

$$T_{(a,b)}(\tau) = \frac{1}{N}\sum_{k\in\mathbb{Z}/N\mathbb{Z}} e\left(\frac{-bk}{N}\right) T(a,k,\tau)$$

allow us to pass from one equation to the other - this is essentially Lemma 3.3 of [Carnahan-2012] but we multiply by the additional character $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right) \mapsto e(-c/24)$, $\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right) \mapsto 1$ of $SL_2(\mathbb{Z})$.

Since $W^{(a(1,\frac{t_n r}{(n,t)})+b(0,\frac{rn}{(n,t)}))} = V^{(a_n, b_n+2t_n\lfloor a/n\rfloor)}$, we have

$$T_{(a(1,\frac{t_n r}{(n,t)})+b(0,\frac{rn}{(n,t)}))}(\tau) = T_{(a_n,b_n+2t_n\lfloor a/n\rfloor)}(\mathbf{1},\tau)$$

for all $a,b\in\mathbb{Z}/N\mathbb{Z}$, and $T_{(a,b)}(\tau) = 0$ for $(a,b)\notin B'$. By our computation of the quadratic form in the first paragraph of this proof, we find that

$$\begin{aligned} T_{(a(1,\frac{t_n r}{(n,t)})+b(0,\frac{rn}{(n,t)}))}(\tau+1) &= T_{(a_n,b_n+2t_n\lfloor a/n\rfloor)}(\mathbf{1},\tau+1) \\ &= e\left(\frac{a_n(b_n+2t_n\lfloor a/n\rfloor)}{n} + \frac{a_n^2 t_n}{n^2} - \frac{c}{24}\right) T_{(a_n,b_n+2t_n\lfloor a/n\rfloor)}(\mathbf{1},\tau) \end{aligned}$$

Then $\frac{a_n(b_n+2t_n\lfloor a/n\rfloor)}{n} + \frac{a_n^2 t_n}{n^2} = \frac{ab}{n} + \frac{a^2 t_n}{n^2} = \frac{abrn+a^2t_nr}{N(n,t)} = \frac{a\left(\frac{at_nr}{(n,t)}+\frac{brn}{(n,t)}\right)}{N}$, so we obtain

$$T_{(a(1,\frac{t_nr}{(n,t)})+b(0,\frac{rn}{(n,t)}))}(\tau+1) = e\left(\frac{a\left(\frac{at_nr}{(n,t)}+\frac{brn}{(n,t)}\right)}{N} - \frac{c}{24}\right) T_{(a(1,\frac{t_nr}{(n,t)})+b(0,\frac{rn}{(n,t)}))}(\tau).$$

This proves the third equation.

For the fourth equation, we note that by Theorem 2.2.8, the S-matrix entries $S_{(i,j),(k,l)}$ for ${}^gV$ are given by $\frac{1}{n}e(b_\Omega((i,j),(k,l)))$, where b_Ω is -1 times the inner product induced by q_D. Since the map $B' \to D$ preserves the quadratic form, hence -1 times the $\mathbb{Q}/\mathbb{Z}$-valued inner product, if we write b'_Ω for the pullback of b_Ω, we have

$$\begin{aligned}
T_{(a(1,\frac{t_nr}{(n,t)})+b(0,\frac{rn}{(n,t)}))}(\frac{-1}{\tau}) &= T_{(a_n,b_n+2t_n\lfloor a/n\rfloor)}(\mathbf{1},\frac{-1}{\tau}) \\
&= \frac{1}{n}\sum_{x\in D} e(b_\Omega((a_n,b_n+2t_n\lfloor a/n\rfloor),x))T_x(\mathbf{1},\tau) \\
&= \frac{1}{N}\sum_{y\in B'} e(b'_\Omega((a,\frac{at_nr+brn}{(n,t)}),y))T_y(\tau) \\
&= \frac{1}{N}\sum_{p,q\in\mathbb{Z}/n\mathbb{Z}} e\left(-\frac{\frac{at_nr+brn}{(n,t)}p+aq}{N}\right)T_{(p,q)}(\tau),
\end{aligned}$$

so the fourth equation holds for $(a,b)\in B'$. For $(a,b)\notin B'$, we need a vanishing result, and it suffices to show that $\sum_{p,q\in\mathbb{Z}/n\mathbb{Z}} e(-\frac{bp+aq}{N})T_{(p,q)}(\tau) = 0$. This sum reduces to a sum over B', and it suffices to show that for any $x\in D$, the sum over preimages of x in B' vanishes. The kernel of $B'\to D$ is generated by $(n,\frac{-nrt_n}{(n,t)}) = n(1,\frac{t_nr}{(n,t)}) - 2t_n(0,\frac{rn}{(n,t)})$, so we fix $p(1,\frac{t_nr}{(n,t)})+q(0,\frac{rn}{(n,t)})\in B'$ and compute the coefficient on $T_{p(1,\frac{t_nr}{(n,t)})+q(0,\frac{rn}{(n,t)})}(\mathbf{1},\tau)$:

$$\begin{aligned}
&\sum_{k=0}^{N/n-1} e\left(\frac{b(p+kn)+a(\frac{pt_nr+qrn-knrt_n}{(n,t)})}{N}\right) \\
&\quad= \sum_{k=0}^{N/n-1} e\left(\frac{bp(n,t)+bkn(n,t)+aprt_n+aqrn+aknrt_n}{N(n,t)}\right) \\
&\quad= e\left(\frac{bp(n,t)+aprt_n+aqrn}{N(n,t)}\right)\sum_{k=0}^{N/n-1} e\left(k\frac{bn(n,t)+anrt_n}{N(n,t)}\right)
\end{aligned}$$

The last sum vanishes if and only if the numerator $bn(n,t)+anrt_n$ is nonzero modulo $N(n,t)$, and by fixing a, a short calculation shows that this is precisely when $(a,b)\notin B'$. ☐

Generalized Moonshine for cyclic groups was proved in [Dong-Li-Mason-2000] up to a scalar ambiguity. We can eliminate the ambiguity by using a central extension, essentially given by unrolling the corresponding abelian intertwining algebra.

Definition 2.3.5. For any $g\in\mathbb{M}$, and let N be any positive integer such that the McKay-Thompson series $T_g(\tau) = \mathrm{Tr}(gq^{L_0-1}|V^\natural)$ is $\Gamma_0(N)$-invariant. By [Carnahan-2012] Corollary 3.25, we may define the vector-valued function $\hat{F}^g = \{\hat{F}^g_{i,j}\}_{i,j=0}^{N-1}$ by $\hat{F}^g_{i,j} = T_{g^{(i,j,N)}}(\tau)|_0 A$ for $A = \begin{pmatrix} * & * \\ \frac{i}{(i,j,N)} & \frac{j}{(i,j,N)}\end{pmatrix} \in SL_2(\mathbb{Z}/\frac{N}{(i,j,N)}\mathbb{Z})$. Similarly, we may define the vector valued function $F^g_{i,j} =$

$\{F^g_{i,j}\}_{i,j=0}^{N-1}$ by the discrete Fourier transform $F^g_{i,j}(\tau) = \frac{1}{N}\sum_{k\in\mathbb{Z}/N\mathbb{Z}} e\left(\frac{-jk}{N}\right)\hat{F}^g_{i,k}(\tau)$. This is a vector-valued modular function for the Weil representation $\rho_{II_{1,1}(1/N)}$. We write $c^g_{i,j}(n)$ for the q^n coefficient of $F^g_{i,j}$.

Corollary 2.3.6. *For any $g \in \mathbb{M}$, hypothesis B_g from section 4.4 of [Carnahan-2012] is satisfied, i.e., if $T_g(\tau)$ has level N, then there are actions $\{\psi_i\}_{i=0}^{N-1}$ of $\mathbb{Z}/N\mathbb{Z}$ on the twisted modules $V^\natural(g^i)$, compatible with the action of $\langle g\rangle$ on $V^\natural$, such that $\mathrm{Tr}(\psi^i(j)q^{L_0-1}|\dim V^\natural(g^i)) = \hat{F}^g_{i,j}(\tau)$, and $\psi_i(i) = e(L_0)$.*

Proof. The actions ψ_i given in Proposition 2.3.4 satisfy these identities, and $V^\natural$ satisfies the hypotheses of that proposition by Theorem 2.2.9. □

2.4. **Adding a spacetime torus.** In order to produce a Lie algebra from a quantization functor, we need a conformal vertex algebra of central charge 26. Our abelian intertwining algebras ${}^gV$ and ${}^g_NV$ have central charge 24, so we have to increase the central charge by two and cancel the braiding. Our functors amount to taking a graded tensor product with the abelian intertwining algebra attached to a well-chosen 2-dimensional hyperbolic lattice.

Definition 2.4.1. Let L be an integer lattice of signature $(1,1)$, and let (D,q) be the discriminant form $L^\vee/L$. We define the "add a torus" functor AT_L from $AIA^c_{(D,q)}$ to the category of L-graded conformal vertex algebras of central charge $c+2$ as follows:

(1) For any object $W = \bigoplus_{a\in D} W_a$ in $AIA^c_{(D,q)}$ with abelian 3-cocycle (F,Ω), we choose an abelian 3-coboundary to endow $V_{L^\vee(-1)}$ with the abelian 3-cocycle $(\frac{1}{F},\frac{1}{\Omega})$, and take the graded tensor product $W\otimes^D V_{L^\vee(-1)} = \bigoplus_{a\in D} W_a \otimes V_{L_{-1}+a}$.
(2) For any isomorphism $W \to W'$ in $AIA^c_{(D,q)}$, the corresponding isomorphism of $L^\vee$-graded conformal vertex algebras is given by taking the tensor product.

For any $\lambda \in L^\vee$, the λ-graded piece is the tensor product $W_{\lambda+L}\otimes\pi^{L\otimes\mathbb{C}}_\lambda$.

Lemma 2.4.2. *Let V be a holomorphic C_2-cofinite vertex operator algebra of central charge c, let g be an automorphism of order n such that $V(g^i)$ has strictly positive L_0-spectrum for $g^i \neq 1$, and let ${}^gV$ be the abelian intertwining algebra given in Theorem 2.2.8. Then there is a lattice L of signature $(1,1)$ and an abelian intertwining algebra $V_{L^\vee(-1)}$ whose abelian 3-cocycle cancels that of ${}^gV$. In particular, adding a torus is well-defined on ${}^gV$, and yields an $L^\vee(-1)$-graded conformal vertex algebra of central charge $c+2$.*

Proof. Let L be the lattice identified in [van Ekeren, et al, 2015] by its Gram matrix $\left(\begin{smallmatrix} -2t_n & n \\ n & 0\end{smallmatrix}\right)$. Then the quadratic form on $L^\vee/L$ is precisely q_Ω for ${}^gV$. Thus, there exists a suitable coboundary for $V_{L^\vee(-1)}$ such that the resulting abelian 3-cocycle cancels that of ${}^gV$. □

Proposition 2.4.3. *Adding a torus 2-commutes with both unrolling against hyperbolic lattice embeddings and twisting against abelian 3-coboundaries. That is, different choices of unrolling or coboundaries on an abelian intertwining algebra yield isomorphic conformal vertex algebras.*

Proof. The claim about coboundaries follows from the fact that if one twists against a coboundary for the object in $AIA^c_{(D,q)}$, the opposite coboundary is chosen on $V_{L^\vee(-1)}$. Thus, the graded tensor product is unchanged.

Suppose we are given an embedding $L \subset M$ of integral lattices of signature $(1,1)$, and let $(A,q_A) = M^\vee/M$ and $(B,q_B) = L^\vee/L$. We have the diagram $M^\vee/M \twoheadleftarrow M^\vee/L \hookrightarrow L^\vee/L$ of morphisms of quadratic groups, so the unrolling functor yields an abelian intertwining algebra on $L^\vee/L$ that is supported on $M^\vee/L$. Adding a torus to the A-graded object yields an $M^\vee(-1)$-graded conformal vertex algebra, while adding a torus to the B-graded object yields an $L^\vee(-1)$-graded conformal vertex algebra that is supported on the sublattice $M^\vee(-1)$. Furthermore, the homogeneous components are isomorphic, and the multiplication laws match, since the multiplication law on the $L^\vee(-1)$-graded vertex algebra is given by a tensor product of a lattice algebra with a multiplication rule pulled back from the A-graded abelian intertwining algebra. $\square$

Definition 2.4.4. For any $g \in \mathbb{M}$, we write ${}^gV^\natural$ for the abelian intertwining algebra in $AIA^{24}_{q_D}$ given by Theorem 2.2.8 and Theorem 2.2.9. We write ${}^g_NV^\natural$ for the "unrolled" abelian intertwining algebra in $AIA^{24}_{q_B}$ given by Proposition 2.3.4.

Recall that by forgetting some multiplicative structure, we may view ${}^gV^\natural$ as a direct sum of irreducible g^i-twisted $V^\natural$-modules, as g^i ranges over the cyclic group generated by g. We note that similarly, ${}^g_NV^\natural$ can be viewed as $\bigoplus_{i=0}^{N-1} V^\natural(g^i)$.

Corollary 2.4.5. *For any $g \in \mathbb{M}$, and any multiple N of the level of $T_g(\tau)$, any choice of lattice embedding $I\!I_{1,1}(N) \hookrightarrow L$ (for $L^\vee/L$ the discriminant form for ${}^gV^\natural$) induces an isomorphism $AT_L({}^gV^\natural) \cong AT_{I\!I_{1,1}(N)}({}^g_NV^\natural)$ of $I\!I_{1,1}(-1/N)$-graded conformal vertex algebras of central charge 26.*

Proof. Theorem 2.2.9 implies Proposition 2.4.3 is applicable, and the result immediately follows. $\square$

2.5. **Projective action of centralizer.** We show that for any finite order automorphism g of a holomorphic C_2-cofinite vertex operator algebra V of CFT type, there is a natural action by a central extension of the centralizer of g by homogeneous automorphisms on the abelian intertwining algebra ${}^gV$, the unrolled object ${}^g_NW$, and the conformal vertex algebra given by adding a torus.

Lemma 2.5.1. *Let V be a holomorphic C_2-cofinite vertex operator algebra of CFT type, and let g be an automorphism of order n such that $V(g^i)$ has strictly positive L_0-spectrum for $g^i \neq 1$. Let $h \in C_{\mathrm{Aut}\,V}(g)$. Then the set of homogeneous automorphisms of the abelian intertwining algebra ${}^gV$ that restrict to h on V has precisely n elements, and a natural simply transitive action of $\mu_n(\mathbb{C})$.*

Proof. If $\tilde{h}$ is a homogeneous automorphism of the abelian intertwining algebra ${}^gV$ that restricts to h on V, then for each $g^i \in \langle g \rangle$, the restriction of $\tilde{h}$ to $V(g^i)$ is an element of $\mathrm{Aut}\,V(g^i)$ over h. Now, suppose we choose arbitrary elements $h_i \in \mathrm{Aut}\,V(g^i)$ over h for all i. By Remark 2.1.5, the set of lifts of h to an automorphism of $V(g^i)$ has a simply transitive action of $\mathbb{C}^\times$, so we can measure the failure of $\{h_i\}_{i=0}^{n-1}$ to form an automorphism of ${}^gV$ by setting $\lambda_{i,j} \in \mathbb{C}^\times$ to be the unique constant such that $Y(h_ia, z)h_jb = \lambda_{i,j}h_{i+j}Y(a, z)b$ for all $a \in V(g^i)$ and $b \in V(g^j)$. By the associativity and commutativity of fusion, $\lambda = \{\lambda_{i,j}\}$ is a normalized symmetric 2-cocycle on $\mathbb{Z}/n\mathbb{Z}$ with values in $\mathbb{C}^\times$, hence it defines an abelian central extension E of $\mathbb{Z}/n\mathbb{Z}$ by $\mathbb{C}^\times$. Because $\mathbb{C}^\times$ is divisible, the defining exact sequence splits, and $E \cong \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{C}^\times$. In particular, the cocycle λ is a coboundary, and the set of 1-cochains in the preimage of λ under the differential is a torsor under the group $\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^\times) \cong \mu_n(\mathbb{C})$ of 1-cocycles. Concretely, this means any normalized symmetric 2-cocycle λ is attainable by a suitable choice of $\{h_i\}$, and those $\{h_i\}$ that yield an automorphism of ${}^gV$ (namely those for which $\lambda = 1$) form a torsor under $\mu_n(\mathbb{C})$, where the action is by rescaling: For each $\zeta \in \mu_n(\mathbb{C})$, one replaces h_i with ζ^ih_i. $\square$

Proposition 2.5.2. *Let V be a holomorphic C_2-cofinite vertex operator algebra of CFT type, and let g be an automorphism of order n such that $V(g^i)$ has strictly positive L_0-spectrum for $g^i \neq 1$. Then there is a central extension $\widetilde{C_{\mathrm{Aut}\,V}(g)}$ of $C_{\mathrm{Aut}\,V}(g)$ by $\mu_n(\mathbb{C})$ with a natural action on the abelian intertwining algebra ${}^gV$, the unrolled object ${}^g_NW$, and the conformal vertex algebra given by adding a torus.*

Proof. By Lemma 2.5.1, the set of all lifts of elements of $C_{\mathrm{Aut}\,V}(g)$ is a $\mu_n(\mathbb{C})$-torsor over $C_{\mathrm{Aut}\,V}(g)$, so in order to prove it is a central extension, we need to show that it is closed under multiplication, the torsor structure map is a homomorphism, and all lifts of identity are central. All lifts of identity are central, because the restriction of a lift of identity to any $V(g^i)$ is a scalar, and all lifts of elements are sums of linear automorphisms of the twisted modules $V(g^i)$. The remaining claims about the multiplicative structure follow from the following calculation: if $\tilde{h} = \{h_i\}$ and $\tilde{k} = \{k_i\}$ are lifts of $h, k \in C_{\operatorname{Aut} V}(g)$, then for all $a \in V(g^i), b \in V(g^j)$,

$$\begin{aligned} h_{i+j}k_{i+j}Y(a,z)b &= h_{i+j}Y(k_i a, z)k_j b \\ &= Y(h_i k_i u, a)h_j k_j v b \end{aligned}$$

Therefore, $\tilde{h}\tilde{k}$ is a lift of hk to ${}^gV$.

We thus obtain a natural action of $\widetilde{C_{\operatorname{Aut} V}}(g)$ on ${}^gV$. Since unrolling and adding a torus are functors, we get natural actions of the same group on the resulting objects. ☐

Lemma 2.5.3. *For any N divisible by the level of $T_g(\tau)$, the $II_{1,1}(1/N)$-homogeneous conformal vertex algebra isomorphism between $AT_L({}^gV^\natural)$ and $AT_{II_{1,1}(N)}({}^g_NV^\natural)$ given in Corollary 2.4.5 can be made $\widetilde{C_{\mathbb{M}}(g)}$-equivariant.*

Proof. Promotion to $\widetilde{C_{\mathbb{M}}(g)}$-equivariance follows from the structure of the unrolling functor - the corresponding graded pieces are identified as $\widetilde{C_{\mathbb{M}}(g)}$-modules, and the multiplication maps are equivariant. ☐

We summarize the results of this section.

Theorem 2.5.4. *Let g be an automorphism of $V^\natural$, i.e., an element of the Monster simple group $\mathbb{M}$, and suppose the McKay-Thompson series $T_g(\tau)$ has level N. Let $B = \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ with the quadratic form $q_B(a,b) = e(ab/N)$. Then there is a central extension $\widetilde{C_{\mathbb{M}}(g)}$ of the centralizer $C_{\mathbb{M}}(g)$ by $\mu_{|g|}(\mathbb{C})$ equipped with a distinguished lift $\tilde{g}$ of g, and an abelian intertwining algebra ${}^g_NV^\natural \in AIA^{24}_{q_B}$ with a homogeneous action of $\widetilde{C_{\mathbb{M}}(g)}$ by autmorphisms, satisfying the following properties:*

1. *${}^g_NV^\natural$ decomposes into $\bigoplus_{a=0}^{N-1} V^\natural(g^a)$ as a direct sum of twisted $V^\natural$-modules.*
2. *The action of $\tilde{g}$ restricted to $V^\natural(g)$ is given by $e(L_0)$.*
3. *The (a,b)-graded component of ${}^g_NV^\natural$ is $\{v \in V^\natural(g^a) | \tilde{g}v = e(b/N)v\}$. This is nonzero if and only if $(a,b) \in B'$ (where $B' \subseteq B$ is a subgroup uniquely determined by the symmetries of T_g), and in that case it is a simple current as a $(V^\natural)^g$-module, and it is endowed with a $\widetilde{C_{\mathbb{M}}(g)}$-action.*
4. *The graded dimension of the (a,b)-graded component ${}^g_NV^\natural$ is given by the component form $F^g_{a,b}(\tau)$ given in Definition 2.3.5, and in particular, the coefficient $c^g_{a,b}(n)$ is equal to the dimension of $\{v \in V^\natural(g^a) | \tilde{g}v = e(b/N)v, L_0 v = (n+1)v\}$. That is, the vector-valued graded dimension of ${}^g_NV^\natural$ is equal to F^g, a vector valued modular function for the Weil representation attached to the lattice $II_{1,1}(N)$.*

The conformal vertex algebra $AT_{II_{1,1}(N)}({}^g_NV^\natural)$ has central charge 26, and is graded by $II_{1,1}(-1/N)$, with an action of $\widetilde{C_{\mathbb{M}}(g)}$ by homogeneous automorphisms.

Proof. As before, Theorem 2.2.9 allows us to use the general results for a holomorphic C_2-cofinite vertex operator algebra V of CFT type equipped with an automorphism g for which $V(g^i)$ has strictly positive L_0-spectrum whenever $g^i \neq 1$.

By Lemma 2.5.1, a lift of g to an automorphism of ${}^g_NV^\natural$ is uniquely determined by what it does to $V^\natural(g)$, and by the $\tau \mapsto \tau + 1$ property given in Proposition 2.3.4, setting $\psi_1 = e(L_0)$ yields a lift $\tilde{g}$ such that the (a,b)-graded component of ${}^g_NV^\natural$ is $\{v \in V^\natural(g^a) | \tilde{g}v = e(b/N)v\}$. The claim that the (a,b)-component is nonzero if and only if $(a,b) \in B'$ is straightforward from the definition of unrolling, and the simple current assertion is given by the first listed claim in Proposition 2.3.4.

The claims about graded dimension follow from the properties of the functions $T(a,b;\tau)$ and $T_{(a,b)}(\tau)$ given in Proposition 2.3.4. In particular, [Carnahan-2012] Lemma 3.6 implies the functions

$\{T(a,b,\tau)\}_{a,b\in\mathbb{Z}/N\mathbb{Z}}$ and $\{F^g_{a,b}(\tau)\}_{a,b\in\mathbb{Z}/N\mathbb{Z}}$ on $\mathcal{M}^{\mathbb{Z}/N\mathbb{Z}}_{Ell}$ are uniquely determined by $T(0,m,\tau)$ and $F^g_{0,m}(\tau)$ for m ranging over divisors of N, and in both cases, these are given by the McKay-Thompson series $T_{g^m}(\tau)$. The discrete Fourier transform then gives an equality between $T_{(a,b)}(\tau)$ and $F^g_{a,b}(\tau)$, matching the coefficient $c^g_{a,b}(n)$ with the dimension of the stated eigenspace.

The remaining claims follow from Proposition 2.5.2 and Corollary 2.4.5. □

3. Lie algebras

In this section we construct the monstrous Lie algebras. Our construction employs the physical principle first articulated in [Lovelace-1971] and demonstrated in [Goddard-Thorn-1972] and [Brower-1972], that in the critical dimension $d = 26$ of bosonic string theory, Virasoro-type Ward identities cancel two full sets of oscillators. This principle now has at least three methods of realization in physics (cf. [Green-Schwarz-Witten-1987] chapters 2,3): Covariant quantization, passage to the light-cone gauge (also called "construction of the transversal space"), and BRST cohomology. The use of the BRST method for quantizing strings was pioneered in [Kato-Ogawa-1983].

The use of this cancellation principle in mathematics is also far from new. For example, BRST cohomology was used in [Frenkel-Garland-Zuckerman-1986] to compute the character of the space of physical states for a bosonic string propagating in Lorentzian 26-space. The covariant quantization method is also used in the analysis of three Lie algebras, each of which was named the "Monster Lie algebra" upon construction:

1. The first example is the Kac-Moody Lie algebra whose simple roots are those of the reflection group of the even unimodular Lorentzian lattice $II_{1,25}$. The existence of a Monster action was suggested in [Borcherds-Conway-Queen-Sloane-1984], and Frenkel [Frenkel-1985] used the no-ghost theorem to bound its root multiplicities by coefficients of the modular form $1/\Delta$.
2. The second example is the Borcherds-Kac-Moody Lie algebra whose real simple roots are those given in the first example, but it has additional imaginary simple roots of multiplicity 24 on positive multiples of the Weyl vector. This was constructed in [Borcherds-1986], and is now known as the "fake Monster Lie algebra". In [Borcherds-1990], Borcherds showed that the root multiplicities saturate the $1/\Delta$ bound given by Frenkel.
3. The third and last example is the generalized Kac-Moody Lie algebra constructed in [Borcherds-1992]. This is the only example known to have a faithful Monster action.

Supersymmetric variants of this construction have been used, e.g., in [Scheithauer-2000].

3.1. **Virasoro preliminaries.** Recall the Virasoro algebra vir is the complex vector space with a distinguished basis given by symbols $K, \{L_n\}_{n\in\mathbb{Z}}$, with the Lie bracket defined by $[L_m, L_n] = (m-n)L_{m+n} + \frac{m^3-m}{12}\delta_{m+n,0}K$ for all $m,n \in \mathbb{Z}$, and $[L_n, K] = 0$ for all $n \in \mathbb{Z}$. We write vir^+ for the subalgebra spanned by $\{L_n\}_{n\geq 1}$, vir^- for the subalgebra spanned by $\{L_n\}_{n\leq -1}$, and vir^0 for the 2-dimensional subalgebra spanned by L_0 and K.

Definition 3.1.1. Let V be a representation of the Virasoro algebra, i.e., a complex vector space equipped with operators $\{L_n\}_{n\in\mathbb{Z}}$ and K that satisfy the Virasoro relations.

1. We say that V is positive-energy if:
   (a) L_0 acts diagonalizably, i.e., V splits into a direct sum of L_0-eigenspaces.
   (b) vir^+ acts locally nilpotently, i.e., for all $v \in V$, there exists $N > 0$ such that for all sequences $n_1, \dots, n_k$ of positive integers satisfying $n_1 + \cdots + n_k > N$, we have $L_{n_1} \cdots L_{n_k} v = 0$.
2. We say that V has central charge $c \in \mathbb{C}$ if $Kv = cv$ for all $v \in V$.
3. A bilinear or sesquilinear form $(,)$ on V is Virasoro-invariant if for each $n \in \mathbb{Z}$, L_n is adjoint to L_{-n}.

(4) We say that V is unitarizable if it admits a positive-definite hermitian Virasoro-invariant form.

(5) A vector $v \in V$ is:
  (a) primary (or physical) if $L_i v = 0$ for all $i > 0$, i.e., if vir^+ acts trivially.
  (b) spurious if it has the form $L_{-i}w$, for some $w \in V$ and some $i > 0$.
  (c) null if it is primary and spurious.

  We write P^r to denote the subspace of primary vectors v such that $L_0 v = rv$. We may write P^r_V when there is more than one representation under consideration. We write $null^r_V$ to denote the subspace of null vectors of weight r in V.

Given $c, h \in \mathbb{C}$, we define the Verma module $M_{h,c}$ as the induced module $vir \otimes_{vir^+ \oplus vir^0} \mathbb{C}_{h,c}$, where $\mathbb{C}_{h,c}$ is the one dimensional vector space with trivial action of vir^+, such that L_0 has eigenvalue h, and K has eigenvalue c. We write $L_{h,c}$ for the unique irreducible quotient of $M_{h,c}$ with highest weight h.

We note that any Verma module $M_{h,c}$ is positive-energy and admits a canonical invariant bilinear form, called the Shapovalov form, which is nonsingular if and only if $M_{h,c}$ is irreducible (see [Feigin-Fuchs-1984] or [Kac-Raina-1987]). In the following lemma, we set $c = 24$, because that is the case we will use. Similar results hold for other central charges in the open interval $(1, 25) \subset \mathbb{R}$.

Lemma 3.1.2. *Let V be a positive energy vir-module of central charge $c = 24$. Then V is unitarizable if and only if the L_0-spectrum of V is non-negative and L_{-1} annihilates the zero weight subspace $V_0 = \operatorname{Ker} L_0$. If these conditions hold, then V decomposes as an orthogonal direct sum of irreducible vir-modules as follows:*

$$V \cong \bigoplus_{h \geq 0} P^h_V \otimes L_{h,24} \cong (V_0 \otimes L_{0,24}) \oplus \bigoplus_{h>0} P^h_V \otimes M(c,h)$$

Proof. The Kac determinant formula yields the following unitarizability and irreducibility results (where detailed proofs can be found in the first four sections of [Kac-Raina-1987]):

(1) If $h \neq 0$, then $M_{h,24}$ is irreducible (i.e., equal to $L_{h,24}$), and the category of lowest weight vir-modules with central charge 24 and lowest weight h is completely reducible, with $L_{h,24}$ as its unique irreducible object.
(2) For the case $h = 0$, $M_{0,24}$ has a unique proper submodule $M_{1,24}$, and $L_{0,24} = M_{0,24}/M_{1,24}$.
(3) $M_{h,24}$ is unitarizable if and only if $h > 0$. $L_{h,24}$ is unitarizable if and only if $h \geq 0$.

Given V unitarizable, we choose a Virasoro-invariant Hermitian form. By self-adjointness of L_0, any two L_0-eigenspaces with different eigenvalues are orthogonal. For any coset $h + \mathbb{Z}$, we let $V_{h+\mathbb{Z}}$ denote the subspace spanned by L_0-eigenvectors with eigenvalues in $h+\mathbb{Z}$. Then any disjoint cosets of $\mathbb{Z}$ yield orthogonal subspaces of V, so we may consider one coset at a time. Within each coset, the decomposition follows from induction by L_0-eigenvalue, noting that within each L_0-eigenspace the primary subspace is orthogonal to the spurious subspace. The necessary and sufficient condition then follows immediately. □

Lemma 3.1.3. *Let g be an automorphism of $V^\natural$. Then the irreducible g-twisted module $V^\natural(g)$ is unitarizable. If $g \neq 1$, then $V^\natural(g)$ decomposes as a direct sum of Virasoro submodules of the form $M(c,h)$ as h ranges over a multiset of positive rational numbers of the form $\frac{k}{|g|^2}$.*

Proof. For the case $g = 1$, it suffices to show that $V^\natural$ is unitarizable - this is well-known, and follows straightforwardly from its conformal structure. By Theorem 2.2.9, if $g \neq 1$, then the lowest weight space of $V^\natural(g)$ has positive L_0-eigenvalue, so by Lemma 3.1.2, we may decompose $V^\natural(g)$ as a sum of unitarizable Verma modules. □

Lemma 3.1.4. *Let $\lambda \in \mathbb{R}^{r,s}$, and let $\pi^{r,s}_\lambda$ denote the corresponding irreducible Heisenberg module, endowed with the Virasoro action arising from the conformal element $\frac{1}{2}\sum_\mu a^\mu_{-1} b^\mu_{-1} 1 \in \pi^{r,s}_0$. Then*

$\pi^{r,s}_\lambda$ is a positive-energy vir-module of central charge 2 that admits a natural Virasoro-invariant Hermitian form for which b^μ_n is adjoint to b^μ_{-n}, for all $n \in \mathbb{Z}$ and all vectors $b^\mu \in \mathbb{R}^{r,s}$.

Proof. By irreducibility, the conditions on the Hermitian form determine it up to a normalization. See, e.g., section 2.2 of [Frenkel-Garland-Zuckerman-1986]. ☐

Lemma 3.1.5. *Let V be a Virasoro representation equipped with a Virasoro-invariant bilinear form $(,)$.*

1. *The L_0-eigenspaces in V are pairwise orthogonal with respect to the form.*
2. *$L_{-1}P^0 \subseteq P^1$. In particular, $L_{-1}P^0 \subseteq null^1_V$.*
3. *A vector in P^1 is orthogonal to all other vectors in P^1 if and only if it lies in $null^1_V$.*

Proof. All claims follow from straightforward calculations. ☐

There is a stronger notion of invariance for vertex algebras (found, e.g., in [Frenkel, et al, 1993]).

Definition 3.1.6. Let V be a conformal vertex algebra. A bilinear form $(,)$ on V is invariant if for all $u, v, w \in V$, the identity $(Y(u,z)v, w) = (v, Y(e^{zL_1}(-z^{-2})^{L_0}u, z^{-1})w)$ holds. Equivalently, if u has conformal weight i, then the adjoint of the operator u_n is $(-1)^i \sum_{j\geq 0}(L_1^j(u))_{2i-j-n-2}/j!$.

Lemma 3.1.7. *Let V be a conformal vertex algebra. Then any invariant bilinear form is Virasoro-invariant.*

Proof. This follows from evaluating the defining identity with u chosen to be the Virasoro vector ω. ☐

Proposition 3.1.8. *Let V be a holomorphic C_2-cofinite vertex operator algebra of CFT type, and let g be a finite order automorphism of V such that the twisted modules $V(g^i)$ have strictly positive L_0-spectrum for all $g^i \neq 1$. Let L be a hyperbolic lattice such that $L^\vee/L$ has the discriminant form of ${}^gV$, let $M \subset L$ be a hyperbolic sublattice of finite index, and consider the conformal vertex algebra W given by applying AT_M to the $M^\vee/M$-graded abelian intertwining algebra given by unrolling ${}^gV$ (by Proposition 2.4.3, this is isomorphic to $AT_L({}^gV)$ viewed as an $M^\vee(-1)$-graded vertex algebra supported on $L^\vee(-1)$). Let $\langle,\rangle_W$ be the bilinear form on W induced by the contragradient pairings between the component modules $V^{(i,j)}$ in ${}^gV$ and the contragradient pairings between the modules $\pi^{1,1}_\lambda$ and $\pi^{1,1}_{-\lambda}$ for the Heisenberg vertex algebra for $\mathbb{R}^{1,1}$, normalized such that*

$$\langle \mathbf{1}, Y(e^{zL_1}(-z^2)^{-L_0}u, z^{-1})v\rangle_W = \langle Y(u,z)\mathbf{1}, v\rangle_W$$

for all $u \in ({}^gV)^\lambda \otimes \pi^{1,1}_\lambda$ and $v \in ({}^gV)^{-\lambda} \otimes \pi^{1,1}_{-\lambda}$. Then $\langle,\rangle_W$ is an invariant bilinear form on W satisfying the following conditions:

1. *For any $i, j \in L^\vee$ satisfying $i + j \neq 0$, we have $\langle u, v\rangle_W = 0$ for any $u \in ({}^gV)^i$ and $v \in ({}^gV)^j$.*
2. *If $L_0u = \underbrace{ru \text{ and}}_{} L_0v = sv$, and $r \neq s$, then $\langle u, v\rangle_W = 0$.*
3. *If $h \in C_{\operatorname{Aut} V}(g)$, then $\langle hu, hv\rangle_W = \langle u, v\rangle_W$.*

Proof. By Theorem 2.13 in [Scheithauer-1998] (which generalizes earlier results like [Li-1994] Theorem 3.1 to the setting we need), if A is a conformal vertex algebra on which L_1 acts locally nilpotently, the dual space $\operatorname{Hom}(A_0/L_1A_1, \mathbb{C})$ is isomorphic to the space of invariant bilinear forms on A, and an isomorphism is given by sending a map f to the form $(u, v) = (f \circ \pi)(u^*_{-1}v)$, where

1. π is the projection from A to the weight zero space A_0.
2. $u^*_n = (-1)^i \sum_{j\geq 0}(\frac{L_1^j}{j!}u)_{2i-j-n-2}$ when $L_0u = iu$.

We consider the unique functional f on W_0 that sends the unit $\mathbf{1}$ to 1, and annihilates all weight zero $a \otimes b$ for $a \in \pi^{1,1}_\lambda$ and $b \in ({}^gV)^\lambda$ when $\lambda \in L^\vee(-1) \setminus \{0\}$. Since f annihilates all vectors in the image of L_1 on the weight one subspace, it corresponds to an invariant bilinear form $(,)_f$. Since $({}^gV)^0 \otimes \pi^{1,1}_0$ has a unique invariant bilinear form up to normalization, the contragradient pairing coincides with $(,)_f$ on this subspace. It is clear that both $\langle,\rangle_W$ and $(,)_f$ satisfy the homogeneity and $\widetilde{C_{\mathrm{Aut}\,V}}(g)$-invariance conditions listed in the statement of the proposition. It remains to show that $(,)_f$ coincides with $\langle,\rangle_W$ in general, and for that it suffices to restrict our view to a pairing between $u \in ({}^gV)^\lambda \otimes \pi^{1,1}_\lambda$ and $v \in ({}^gV)^{-\lambda} \otimes \pi^{1,1}_{-\lambda}$. In this case,

$$\begin{aligned}(u,v)_f &= (f \circ \pi)(u^*_{-1}v) \\ &= (\mathbf{1}, u^*_{-1}v)_f \\ &= \langle \mathbf{1}, u^*_{-1}v\rangle_W\end{aligned}$$

By our normalization, we have the equality

$$\langle \mathbf{1}, Y(e^{zL_1}(-z^2)^{-L_0}u, z^{-1})v\rangle_W = \langle Y(u,z)\mathbf{1}, v\rangle_W$$

of elements of $\mathbb{C}[z]$. Isolating the constant term yields $\langle \mathbf{1}, u^*_{-1}v\rangle_W = \langle u,v\rangle_W$. We conclude that $(u,v)_f = \langle u,v\rangle_W$, so $\langle,\rangle_W$ is invariant.

□

3.2. **Old covariant quantization.** In covariant quantization, one begins with a naïve space of states, that has redundancies due to internal symmetry. In the case of internal Virasoro symmetry, to obtain the relevant space of states that couple, one considers the "physical" subspace of states that are invariant under the action of vir^+, and quotients by the "spurious" subspace, which is spanned by the image of the augmentation ideal in $U(vir^-)$ together with the radical of the induced Hermitian form.

An enhancement of the original covariant approach was used in [Borcherds-1992] to construct the Monster Lie algebra as the space of physical states arising from the vertex algebra $V^\natural \otimes V_{II_{1,1}(-1)}$. We will use a variant of this method to construct our monstrous Lie algebras.

Definition 3.2.1. Let $\mathrm{Rep}^+_c(vir)$ denote the groupoid whose objects are positive energy Virasoro representations of central charge c with Virasoro-invariant bilinear forms, and whose morphisms are Virasoro-equivariant isometries. We define the functor OCQ, from $\mathrm{Rep}^+_c(vir)$ to the groupoid of vector spaces with bilinear forms, as $V \mapsto P^1_V/null^1_V$.

Lemma 3.2.2. *We list some relatively straightforward facts:*

1. *If W is a conformal vertex algebra that is positive-energy, then $P^1/L_{-1}P^0$ has a Lie algebra structure induced by setting $[u,v] = u_0v$, i.e., the coefficient of z^{-1} in $Y(u,z)v$. If a group G acts on W by conformal vertex algebra automorphisms, then G acts on $P^1/L_{-1}P^0$ by Lie algebra automorphisms.*
2. *If W is a positive-energy conformal vertex algebra equipped with an invariant bilinear form (in particular, with L_i adjoint to L_{-i} for each $i \in \mathbb{Z}$), then the Lie algebra $P^1/L_{-1}P^0$ has an invariant inner product induced from W, and the radical of that inner product is an ideal. In particular, $OCQ(W) = (P^1/L_{-1}P^0)/rad(,)$ is a Lie algebra equipped with a nondegenerate invariant inner product. If a group G acts on W by conformal vertex algebra automorphisms, and the invariant bilinear form is G-invariant, then $OCQ(W)$ has a natural G action by Lie algebra automorphisms, and the induced bilinear form is G-invariant.*

Proof. The first sentence of the first claim is given in section 3 of [Borcherds-1992]. The Lie algebra structure of $W/L_{-1}W$ is known from [Borcherds-1986], so it suffices to show that a_0b is primary if a and b are. This follows from the formula $L_n(a_0b) = a_0L_nb + \sum_{m\geq -1}\binom{n+1}{m+1}(L_ma)_{n-m}b$ for $n \geq -1$ together with the L_{-1}-derivative rule. The second sentence follows from the fact that automorphisms preserve the Virasoro action (hence the property of being primary, the grading, and the L_{-1} operator), and the Lie bracket $(u,v) \mapsto u_0v$.

The first sentence of the second claim is given in the proof of Theorem 6.1 of [Borcherds-1992]. Invariance follows immediately from the adjunction formula $u^*_n = (-1)^i \sum_{j\geq 0}(\frac{L_1^j}{j!}u)_{2i-j-n-2}$ when

$n = 0$, $i = 1$, and $L_1 u = 0$. The ideal property follows immediately from invariance. The part about group actions follows from the same reason as in the first claim. ☐

The following proposition seems to capture the core of Lovelace's principle of oscillator cancellation. It is sometimes given the name "No Ghost Theorem", because it immediately implies that quantization yields a positive-definite space of states. That is, if we choose a unitary structure on V, then transport of structure along the component isomorphisms yields a unitary structure on the output. Here, the word "ghost" means negative-norm state, whose absence is important in physics, since the existence of such ghosts obstructs reasonable formulations of causality. In the next section, we will consider a different type of ghost, named after Faddeev-Popov, and they get their name from their violation of the spin-statistics theorem.

Proposition 3.2.3. *Let V be a unitarizable Virasoro representation of central charge 24 equipped with a non-degenerate Virasoro-invariant bilinear form, let $\pi^{1,1}_\alpha$ be the Heisenberg module attached to the linear function α on $\mathbb{R}^{1,1}$, and let $W = V \otimes \pi^{1,1}_\alpha$. Then:*

1. *If $\alpha \neq 0$, then $OCQ(W)$ is isomorphic to the subspace $V_{1-(\alpha,\alpha)}$ of V on which L_0 acts by $1-(\alpha,\alpha)$, where $(,)$ is the induced bilinear form on the dual space of $\mathbb{R}^{1,1}$.*
2. *If $\alpha = 0$, then $OCQ(W)$ is isomorphic to $V_0 \otimes (\pi^{1,1}_0)_1 \oplus V_1 \otimes (\pi^{1,1}_0)_0 \cong V_0 \oplus V_0 \oplus V_1$, where V_1 is the subspace of V on which L_0 acts by identity.*

Proof. One may find a full proof of a slightly altered version of this claim in the appendix of [Jurisich-1998]. It is an expansion of the sketch of a proof for Theorem 5.1 in [Borcherds-1992], which is in turn a minor alteration of the argument in [Goddard-Thorn-1972]. In order to pass from the proof in [Jurisich-1998] to a proof of our claim, we need to make the following changes:

1. The vector α here is r there, and we remove the restriction that it lie in the lattice $II_{1,1}$, since that restriction is never used in the proof.
2. We remove the unnecessary assumption that V is a vertex operator algebra, as the proof does not use any extra structure on the vector space V except the Virasoro action.
3. The space W here is $\mathcal{H}$ there.
4. Although it isn't necessary for this paper, we remove the assumption that the weight 0 subspace V_0 of V, defined as the kernel of L_0, is one dimensional. One should therefore ignore the paragraph concerning the case $r = 0$ at the end of the proof in [Jurisich-1998]. Instead, we note that all weight one elements of $V \otimes \pi^{1,1}_0$ are primary, so $P^1 \cong V_0 \oplus V_0 \oplus V_1$. The fact that L_{-1} annihilates P^0 follows from the classification of unitarizable Virasoro representations of central charge 24, given in Lemma 3.1.2. By our assumption on V and the fact that $(\pi^{1,1}_0)_1$ has a natural real structure identified with $\mathbb{R}^{1,1}$, the bilinear form on the quotient $P^1_W / L_{-1} P^0_W$ is nondegenerate, i.e., the radical vanishes.

☐

Definition 3.2.4. For each $g \in \mathbb{M}$, we define:

$$\mathfrak{m}_g = OCQ \circ AT_{II_{1,1}(N)}({}^g_N V^\natural)$$

where N is the level of $T_g(\tau)$. This is a Lie algebra, by Lemma 3.2.2, and we call it the monstrous Lie algebra attached to g.

Remark 3.2.5. By Proposition 2.4.3, $\mathfrak{m}_g$ is isomorphic to $OCQ \circ AT_L({}^g V^\natural)$ for any L such that $L^\vee / L$ is the discriminant form for ${}^g V^\natural$. Furthermore, we may replace N with any multiple rN, and get an isomorphic object (although it is now a $II_{1,1}(-1/rN)$-graded Lie algebra supported on the sublattice $II_{1,1}(-1/N)$).

Proposition 3.2.6. *Let $g \in \mathbb{M}$. Then the Lie algebra $\mathfrak{m}_g$ satisfies the following properties:*

(1) $\mathfrak{m}_g$ *is graded by* $II_{1,1}(-1/N)$*, and its degree (0,0) subspace is* $\mathbb{C}^2$*, identified with the weight one space* $(\pi_0^{1,1})_1$ *in the Heisenberg module. Furthermore, the degree is given by the eigenvalue under the adjoint action of the degree* $(0,0)$*-subspace.*

(2) $\mathfrak{m}_g$ *has a nondegenerate invariant bilinear form, and if* u *and* v *are homogeneous vectors of degree* $a, b \in II_{1,1}(-1/N)$ *such that* $a + b \neq 0$*, then* u *and* v *are orthogonal.*

(3) $\mathfrak{m}_g$ *has a canonical action of* $\widetilde{C_{\mathbb{M}}(g)}$ *by homogeneous Lie algebra automorphisms that preserve the bilinear form.*

(4) *For each nonzero* $(a, b) \in II_{1,1}(-1/N)$*, the degree* (a, b) *subspace is isomorphic as a* $\widetilde{C_{\mathbb{M}}(g)}$*-module to the subspace of* $V^\natural(g^a)$ *on which* $\tilde{g}$ *acts by* $e(\frac{b}{N})$ *and* L_0 *acts by* $1 + \frac{ab}{N}$*. The dimension of this subspace is equal to the coefficient* $c^g_{a,b}(ab/N)$ *of the vector-valued modular function* $F^g(\tau)$ *given in Definition 2.3.5. In particular, we have a graded vector space isomorphism between* $\mathfrak{m}_g$ *and the Lie algebra* L_g *defined in [Carnahan-2012] section 4.2 (written* W_g *in Proposition 4.4 of that paper).*

Proof. We prove the claims in the order they are stated.

(1) The grading property follows from the fact that $AT_{II_{1,1}(N)}({}^g_N V^\natural)$ is graded by $II_{1,1}(-1/N)$. Because $({}^gV^0)_1 = 0$ and $({}^gV^0)_0 = \mathbb{C}\mathbf{1}$, Proposition 3.2.3 identifies the degree $(0,0)$ space with $V_0 \otimes (\pi_0^{1,1})_1 = (\pi_0^{1,1})_1$. The eigenvalue assertion follows from the identification of the degree $(0,0)$ subalgebra, and the fact that for $u \in (\pi_0^{1,1})_1$ and $v \in \pi_\alpha^{1,1}$, the z^{-1}-coefficient of $Y(u,z)v$ is $\alpha(u)v$.

(2) The fact that the form is nondegenerate and invariant is given in Lemma 3.2.2. The degree restriction follows from the corresponding property of $AT_{II_{1,1}(N)}({}^g_N V^\natural)$, given in Proposition 3.1.8.

(3) By Proposition 2.5.2 (and Theorem 2.2.9), there is a canonical homogeneous action of $\widetilde{C_{\mathbb{M}}(g)}$ on ${}^gV^\natural$ by homogeneous automorphisms, and this endows $AT_{II_{1,1}(N)}({}^g_N V^\natural)$ with a corresponding action. By the third claim of Lemma 3.2.2, homogeneous automorphisms of $AT_{II_{1,1}(N)}({}^g_N V^\natural)$ are taken to homogeneous Lie algebra automorphisms of $\mathfrak{m}_g$.

(4) The identification of $\widetilde{C_{\mathbb{M}}(g)}$-modules follows from Proposition 3.2.3. By the third listed claim of Theorem 2.5.4, the dimension of this space is equal to the Fourier coefficient $c^g_{a,b}(ab/N)$ of F^g. To compare with the graded parts of L_g, we note that by [Carnahan-2012] Proposition 4.4, the $(a, b/N)$ root space of L_g also has dimension equal to the Fourier coefficient $c^g_{a,b}(ab/N)$ of the vector-valued modular form F^g.

□

3.3. **BRST cohomology.** In addition to covariant quantization, there is the "modern" method of Brecchi, Rouet, and Stora (with Tyutin's independent unpublished work often credited), where one extracts the relevant space of states as a cohomology group. This has the advantage of being "very functorial", but the disadvantage of being rather complicated - one extracts the BRST current and the resulting differential by way of a super-variational principle. In the string-theoretic context, BRST quantization was pioneered in [Kato-Ogawa-1983]. One may find string-theoretic expositions of this method in chapter 3 of [Green-Schwarz-Witten-1987], Chapter 4 of [Polchinski-1998], and section 7 of [D'Hoker-1997], the last of which is intended for a more mathematical audience.

In the mathematical world, BRST quantization of Virasoro representations was introduced using semi-infinite cohomology in [Feigin-1984], and generalized in [Frenkel-Garland-Zuckerman-1986]. Proofs can be found in [Lian-Zuckerman-1991]. Borcherds's use of the no-ghost theorem for moonshine was reinterpreted in terms of BRST cohomology in [Lian-Zuckerman-1995]. We will give a modified presentation that is applicable to the vertex algebra objects we have constructed in the previous section.

We introduce the ghost vertex superalgebra V_{ghost} of central charge -26. By boson-fermion correspondence, for any half-integer λ, there is an isomorphism between the bosonic vertex superalgebra attached to the odd positive definite lattice $\mathbb{Z}$ with conformal vector shifted by λ, and the ghost bc-system of weight $\frac{1}{2}+\lambda$. With the notation of [Frenkel, et al, 1988], the bc system of weight $\frac{1}{2}+\lambda$ is generated by the "anti-ghost" field $b(z)$ that is identified with the bosonic lattice field $z^{-\lambda}X(1,z)$ of conformal weight $\frac{1}{2}-\lambda$, and the "ghost" field $c(z)$ identified with $z^{\lambda}X(-1,z)$ of conformal weight $\frac{1}{2}+\lambda$. The conformal vector is then given by $\frac{1}{2}h(-1)^2+\lambda h(-2)$, and it has central charge $1-12\lambda^2$. For V_{ghost}, we set $\lambda=\frac{3}{2}$.

We briefly describe some cohomology functors and some key properties. There is nothing original in this subsection except perhaps the method of presentation.

Definition 3.3.1. Given a Virasoro ("matter") representation V^m of central charge D and energy-momentum tensor $T^m(z)=\sum_n L^m(n)z^{-n-2}$, we define the Virasoro module $C^{\infty/2+*}(vir,V^m)=V^m\otimes V_{ghost}$, which has central charge $D-26$. The old and new BRST currents are defined to be $j^{BRST,old}=cT^m+\frac{1}{2}:cT^{bc}:$ and $j^{BRST,new}=j^{BRST,old}+\frac{3}{2}\partial^2 c$. The $U(1)_{bc}$ current is $j^{bc}=:bc:$. We consider the following operations on $C^{\infty/2+*}(vir,V^m)$:

(1) The old and new BRST operators

$$\begin{aligned} Q^{old} &= j_0^{BRST,old}=\frac{1}{2\pi i}\oint dz j^{BRST,old}(z)\\ &=\sum_n L^m(n)c_{-n}+\sum_{m<n}(m-n):b_{m+n}c_{-m}c_{-n}:\\ Q^{new} &= j_0^{BRST,new}=\frac{1}{2\pi i}\oint dz j^{BRST,new}(z)\\ &=\sum_{r\in\mathbb{Z}}(L_r^{(m)}-\delta_{r,0})c_{-r}-\frac{1}{2}\sum_{r,s\in\mathbb{Z}}(r-s):c_{-r}c_{-s}b_{r+s}: \end{aligned}$$

(2) The old and new ghost number operators

$$\begin{aligned} U^{old} &= j_0^{bc}=\frac{1}{2}(c_0b_0-b_0c_0)+\sum_{n=1}^{\infty}(c_{-n}b_n-b_{-n}c_n)\\ U^{new} &= U^{old}+\frac{3}{2}=1+c_0b_0+\sum_{n=1}^{\infty}(c_{-n}b_n-b_{-n}c_n) \end{aligned}$$

An eigenvector of U with eigenvalue k is said to have ghost number k.

When $Q^2=0$, we call $C^{\infty/2+*}(vir,V^m)$ the "BRST complex". We also define the "relative subcomplex" ([Zuckerman-1989], section 4) to be the subspace of $V\otimes V_{ghost}$ annihilated by b_0.

Remark 3.3.2. We have "old" and "new" versions of Q and U, to reflect the use in the literature. Papers from the 1980s and early 1990s (e.g., [Kato-Ogawa-1983], [Frenkel-Garland-Zuckerman-1986], [Green-Schwarz-Witten-1987], [Lian-Zuckerman-1991]) use the old versions, while the sources starting from the mid-1990s (e.g., [Polchinski-1998], [D'Hoker-1997], [Lian-Zuckerman-1995]) add an additional $\frac{3}{2}\partial^2 c$ to the BRST current and a constant $3/2$ to the ghost number. This change impacts the output of various functors in the following ways:

(1) The addition of the derivative term makes the new BRST current a "$(1,0)$-form as a quantum operator". Concretely, this means the OPE with the total stress-energy tensor T^m+T^{bc} has the form $T(z)j^{BRST}(0)\sim\frac{1}{z^2}j^{BRST}(0)+\frac{1}{z}\partial j^{BRST}(0)$. In the absence of the derivative, one has a pole of order 3 as the leading term. The extra simplicity makes many calculations easier, especially when quantizing in higher genus settings. However, it is irrelevant for the case at hand, since we just want to compute the cohomology spaces.

(2) When we work out the space of physical states, we find that it lies in ghost degree $-1/2$ in the old convention and 1 in the new convention. In particular, the space of interest was written $H^{-1/2}_{BRST}$ in older work, but is written H^1_{BRST} now. The older convention has the convenient property that Poincaré duality gives a symmetry around zero (see [Frenkel-Garland-Zuckerman-1986] Theorem 1.6 and the remark before Theorem 2.5), while the newer convention is more manageable when considering Gerstenhaber or BV structures on cohomology ([Lian-Zuckerman-1993]).

The calculations relevant to us are unchanged, because a total derivative integrates to zero when evaluating charges. That is, the old and new BRST cohomology functors are naturally isomorphic. To avoid unnecessary confusion, we will write H^1_{BRST} even when referring to results that use the old convention.

Lemma 3.3.3. *Let V^m be a positive-energy representation of Virasoro (called the "matter representation" in the physics literature). We have the following operator calculations in $V^m \otimes V_{ghost}$:*

1. *$Q^2 = 0$ if and only if $D = 26$. That is, $C^{\infty/2+*}(V^m)$ is a complex with differential Q if and only if V^m has central charge 26.*
2. *$[U, Q] = Q$, so Q increases ghost number by one.*
3. *$[U, L_0] = 0$, so ghost number and L_0-eigenvalue provide a bigrading.*

Now, suppose V^m has central charge $D = 26$, so Q gives $V^m \otimes V^{ghost}$ the structure of a complex of vector spaces.

1. *$[Q, L_0] = 0$, so the BRST complex and cohomology are graded by L_0-eigenvalues.*
2. *$[Q, b_0] = L_0$. This implies the relative subcomplex is in fact a subcomplex, and all BRST-closed states with nonzero L_0-eigenvalue are BRST-exact. In particular, the cohomology is supported in weight zero.*

Proof. These claims follow from standard OPE calculations. One may find them enumerated in section 4 of [Zuckerman-1989].

□

Lemma 3.3.4. *If V is a conformal vertex algebra of central charge 26, then $H^1_{BRST}(V)$ is a Lie algebra, and morphisms of conformal vertex algebras are taken to maps of Lie algebras under this functor. In particular, if a group G acts on V by conformal vertex algebra automorphisms, then G naturally acts on $H^1_{BRST}(V)$ by Lie algebra automorphisms.*

Proof. For the Lie algebra claim, see [Lian-Zuckerman-1993] Theorem 2.2. While the authors do not make the explicit claim that morphisms of conformal vertex algebras are taken to Lie algebra homomorphisms, the statement follows from the explicit formulas in their proof together with the faithfulness of the "underlying vector space" functors. The claim about group actions then follows from the functoriality. □

We now have the BRST version of oscillator cancellation.

Proposition 3.3.5. *Let $\alpha \in \mathbb{R}^{1,1}$, and let $\pi^{1,1}_\alpha$ be the Heisenberg representation attached to α. If V is a unitarizable Virasoro representation of central charge 24, then $H^1_{BRST}(V \otimes \pi^{1,1}_\alpha)$ is isomorphic to the subspace $V_{1-\alpha^2}$ of V on which L_0 acts with eigenvalue $1 - \alpha^2$ (if $\alpha \neq 0$), and $V_0 \oplus V_0 \oplus V_1$ (if $\alpha = 0$). In particular, any positive definite Hermitian form on V induces a positive definite Hermitian form on $H^1_{BRST}(V \otimes \pi^{1,1}_\alpha)$.*

Proof. The case $\alpha \neq 0$ is covered in [Zuckerman-1989] Theorem 4.9 under the following translation of notation:

1. $\pi^{1,1}_\alpha$ is written $\mathcal{H}(D, p)$, with $D = 2$ and $p = \alpha$.
2. V is written $\mathcal{K}$.

(3) The BRST differential on $\mathcal{H}(D,p) \otimes \mathcal{K} \otimes V^{gh}$ is written Q_{mod}.
(4) $H^{u+3/2}_{BRST}(V \otimes \pi^{1,1}_\alpha)$ is written $H^u_{Q_{mod}}(p)$ for the same values of u, and $p = \alpha$.
(5) The character $Tr(q^{L_0}|V)$ of $\mathcal{K}$ is written $\chi(\mathcal{K},q) = trq^{J_0}$.

The conclusion of the theorem is that $H^{-1/2}_{Q_{mod}}(p) \cong H^0_{rel}(p, Q_{mod})$, and that $\dim H^0_{rel}(p, Q_{mod})$ is the coefficient of $q^{-(p,p)/2}$ in the shifted character $Tr(q^{L_0-1}|V)$. (The statement given in the paper contains an erroneous division by Δ rather than q - this can be checked by comparison with the free boson case in Theorem 3.4, or using Corollary 2.27 in [Lian-Zuckerman-1991], which is proved in more detail.) This completes the case $\alpha \neq 0$.

The case $\alpha = 0$ follows from an explicit computation of cocycle representatives, following the example of Proposition 2.9 in [Frenkel-Garland-Zuckerman-1986]: One obtains a basis of BRST cohomology given by $\{x \otimes L'_{-1} \wedge L'_{-2} \wedge \cdots\}$ as x ranges over a basis of $(\pi^{1,1}_0 \otimes V)_1 = ((\pi^{1,1}_0)_1 \otimes V_0) \oplus ((\pi^{1,1}_0)_0 \otimes V_1) \cong V_0 \oplus V_0 \oplus V_1$.

As we remarked shortly before the statement of Proposition 3.2.3, the positive definite Hermitian form on $H^1_{BRST}(V \otimes \pi^{1,1}_\alpha)$ is inherited from V by the oscillator cancellation.

□

We conclude with a natural isomorphism between the two quantization functors. Let $\mathbf{1} \in V_{ghost}$ be the ghost vacuum vector - it is the unique (up to scaling) vector that is annihilated by $\{b_n\}_{n\geq 0}$ and $\{c_n\}_{n>0}$, and it has ghost number one.

Lemma 3.3.6. *The map $W \hookrightarrow W \otimes V_{ghost}$ given by $w \mapsto w \otimes \mathbf{1}$ induces a natural isomorphism $OCQ \Rightarrow H^1_{BRST}$, as the input W ranges over* $\mathrm{Rep}^+_{26}(vir)$.

Proof. This is proved at the end of section 4.4 in [Polchinski-1998]. Partial arguments can be found in section 3.2.1 of [Green-Schwarz-Witten-1987] and section 4.6 of [D'Hoker-1997] (which is in lecture 7 part F on the IAS web site).

When W is restricted to be a direct sum of representations of the form $V \otimes \pi^{r,1}_\lambda$ for V unitarizable of central charge $25 - r$ and $r \leq 25$, one may instead appeal to the no-ghost theorem together with complete reducibility in the V part. This is the argument of Lemma 6.7 in [Lian-Zuckerman-1995].

□

3.4. **Borcherds-Kac-Moody property.**

We discuss a class of infinite dimensional Lie algebras defined by Borcherds in [Borcherds-1988], where he called them "generalized Kac-Moody algebras". Borcherds-Kac-Moody algebras are quite similar to Kac-Moody Lie algebras, since they are defined by generators and relations encoded by a generalized Cartan matrix, and they admit representation-theoretic data like Weyl character formulas. They differ in the sense that the conditions defining a suitable generalized Cartan matrix are weaker, and most importantly, simple roots are allowed to be imaginary. We will show that the Monstrous Lie algebras defined earlier in this section are Borcherds-Kac-Moody algebras, using a minor variant of the characterization from [Borcherds-1995], Theorem 1.

The definitions of Borcherds-Kac-Moody algebras (or generalized Kac-Moody algebras) are not uniform in the literature. We will use the definition given in [Jurisich-1996] and [Jurisich-2004], because that is the definition that is both general enough to account for the Lie algebras we need and for which there is a published proof of a homological description of the denominator formula. We needed the homological description in order to produce twisted denominator identities in [Carnahan-2012].

Definition 3.4.1. If I is a countable index set, a matrix $A = (a_{i,j})_{i,j\in I}$ of real numbers is called a generalized Cartan matrix if it satisfies the following conditions:

(1) A is symmetrizable, i.e., there is a diagonal matrix Q whose diagonal entries $Q_{i,i} = q_i$ are positive real numbers, such that QA is symmetric.

(2) $a_{i,j} < 0$ if $i \neq j$.
(3) If $a_{i,i} > 0$, then $a_{i,i} = 2$ and for all $j \in I$, $a_{i,j} \in \mathbb{Z}$.

Given a generalized Cartan matrix $(a_{i,j})_{i,j\in I}$, its universal Borcherds-Kac-Moody algebra $\mathfrak{g}(A)$ is the Lie algebra with generators $\{h_i, e_i, f_i\}_{i,j\in I}$, and the following relations:

1. $\mathfrak{sl}_2$ relations: $[h_i, e_k] = a_{i,k}e_k$, $[h_i, f_k] = -a_{i,k}f_k$, $[e_i, f_j] = \delta_{i,j}h_i$.
2. Serre relations: If $a_{i,i} > 0$, then $\mathrm{ad}(e_i)^{1-2a_{i,j}}(e_j) = \mathrm{ad}(f_i)^{1-2a_{i,j}}(f_j) = 0$.
3. Orthogonality: If $a_{i,j} = 0$, then $[e_i, e_j] = [f_i, f_j] = 0$.

A Borcherds-Kac-Moody algebra is a Lie algebra of the form $(\mathfrak{g}(A)/C).D$, where C is a central ideal, and D is a commutative Lie algebra of outer derivations.

From these starting data, one can build up a theory that closely mirrors that of finite dimensional semisimple Lie algebras, but with some complications. The general theory is substantially developed with proofs in [Jurisich-1996].

There are two main differences between this definition and others:

1. Other definitions may start with the symmetrized matrix QA instead of A. This means one often sees the condition that if $a_{i,i} > 0$, then for all $j \in I$, $2a_{i,j}/a_{i,i} \in \mathbb{Z}$.
2. Other definitions may have central generators $h_{i,j}$ for $i \neq j$, with the relation $[e_i, f_j] = \delta_{i,j}h_i$ replaced with $[e_i, f_j] = h_{i,j}$. As we will see in the next lemma, these necessarily vanish in the cases of interest to us. That is, if we were to use the more general definitions, these generators would end up in the central ideal C.

Most of the Borcherds-Kac-Moody algebras that have naturally appeared in mathematical practice have infinitely many simple roots, but a finite dimensional Cartan subalgebra, i.e., I is infinite, but one chooses C to have finite codimension in the span of $\{h_i\}_{i\in I}$. We give a method to recognize most of the interesting cases:

Lemma 3.4.2. *Any complex Lie algebra $\mathfrak{g}$ satisfying the following conditions is Borcherds-Kac-Moody:*

1. *$\mathfrak{g}$ admits a nonsingular invariant symmetric bilinear form, i.e., $(x, y) = (y, x)$, $(x, [y, z]) = ([x, y], z)$ for all $x, y, z \in \mathfrak{g}$, and if $(x, y) = 0$ for all $y \in \mathfrak{g}$, then $x = 0$.*
2. *$\mathfrak{g}$ has a self-centralizing subalgebra H, called a Cartan subalgebra, such that $\mathfrak{g}$ is the sum of eigenspaces under the action of H, and the nonzero eigenvalues, called roots, have finite multiplicity.*
3. *H has a totally real structure $H_{\mathbb{R}} \subset H$ (i.e., such that $H = H_{\mathbb{R}} \oplus \sqrt{-1}H_{\mathbb{R}}$), such that the restriction of the bilinear form is real-valued, and all roots are elements of the dual $H_{\mathbb{R}}^{\vee}$.*
4. *There exists an element $h \in H_{\mathbb{R}}$, such that:*
    - (a) *the centralizer of h in $\mathfrak{g}$ is H,*
    - (b) *for any $M \in \mathbb{R}$, there exist at most finitely many roots α such that $|\alpha(h)| < M$.*

    This vector h is called a regular element. If $\alpha(h) < 0$ then we say that α is negative, and if $\alpha(h) > 0$ then we say that α is positive.
5. *The norms of roots under the inner product $(,)$ are bounded above.*
6. *Any two roots of non-positive norm that are both positive or both negative have inner product at most zero, and if the inner product is zero, then their root spaces commute.*

Proof. This is essentially theorem 1 in [Borcherds-1995], but we remove the hypothesis that $\mathfrak{g}$ is defined over the real numbers. Borcherds's proof still works, because the reconstruction of $\mathfrak{g}$ by recursively adding simple roots following increasing values of $\alpha(h)$ only uses the real structure on H, not on $\mathfrak{g}$. To match Borcherds's definition with ours, we would need to allow the generators $h_{i,j}$, but Borcherds shows in the course of his proof that $h_{i,j} = [e_i, f_j] = 0$ when $i \neq j$. Thus, the proof shows that we obtain a Borcherds-Kac-Moody algebra in our restricted sense. $\square$

Recall that an element $g \in \mathbb{M}$ is Fricke if and only if the McKay-Thompson series T_g is invariant under the level N Fricke involution $\tau \mapsto \frac{-1}{N\tau}$ for some N.

Proposition 3.4.3. *For each Fricke element $g \in \mathbb{M}$, the Lie algebra $\mathfrak{m}_g$ from Definition 3.2.4 is a Borcherds-Kac-Moody Lie algebra.*

Proof. It suffices to check that the conditions in Lemma 3.4.2 hold.

1. The bilinear form is induced from the contragradient form, as shown in Proposition 3.1.8.
2. We define H to be the $(0,0)$-graded piece of $\mathfrak{m}_g$. By the first claim in Proposition 3.2.6, this is isomorphic to $\mathbb{C}^2$. Furthermore, the left multiplication by H is given by the Heisenberg action on the tensor product vertex algebra, and their eigenvalues are given by degree. The fact that anything not in H has nonzero eigenvalue implies the self-centralizing condition is satisfied.
3. H is naturally identified with $II_{1,1}(-1/N) \otimes_{\mathbb{Z}} \mathbb{C}$, which has a natural real structure $H_{\mathbb{R}} = II_{1,1}(-1/N) \otimes_{\mathbb{Z}} \mathbb{R}$. The inner product on $H_{\mathbb{R}}$ is real valued, and the roots, viewed as elements of the lattice, lie in $H_{\mathbb{R}}^{\vee}$.
4. By identifying $H_{\mathbb{R}}$ with $II_{1,1}(-1/N) \otimes_{\mathbb{Z}} \mathbb{R}$, we let h be any negative norm element whose inner product with any root is nonzero. A standard choice is $(1,2)$, under the identification $II_{1,1}(-1/N) \cong \mathbb{Z} \oplus \frac{1}{N}\mathbb{Z} \subset \mathbb{R}^{1,1}$.
5. If we write the norm of a root at $r \in II_{1,1}(-1/N)$ as r^2, then the fourth claim of Proposition 3.2.6 implies the root space at r is identified with a subspace of ${}^{g}_{N}V^{\natural}$ whose vectors have L_0-eigenvalue equal to $1 - r^2$. The L_0-eigenvalues in ${}^{g}_{N}V^{\natural}$ are bounded below by zero, so the norms of roots are bounded above by 1.
6. Because H has a real Lorentzian structure, we can identify the real span of the root space with $\mathbb{R}^{1,1}$, and any pair of imaginary positive (resp. negative) roots lies in a single quadrant, where their inner products are bounded above by zero. Two such roots are orthogonal if and only if they are positive multiples of the same norm zero vector. We claim that because g is Fricke, there are no positive roots of norm zero. Such a root would have the form $(a,0)$ or $(0,b/N)$ in $II_{1,1}(-1/N)$ for $a,b > 0$, and by the last claim in Proposition 3.2.6, the multiplicities of those root spaces are given by the coefficients $c^g_{a,0}(0)$ and $c^g_{0,b}(0)$ of F^g. By Corollary 3.25 of [Carnahan-2012], these coefficients are in turn the exponents attached to $(1 - p^a q^0)$ and $(1 - p^0 q^{b/N})$ in the product expansion
$$T_g(\sigma) - T_g(-1/\tau) = p^{-1} \prod_{m>0, n \in \frac{1}{N}\mathbb{Z}} (1 - p^m q^n)^{c^g_{m,Nn}(mn)}.$$
Because g is Fricke, we may rewrite the left side as $T_g(\sigma) - T_g(\tau/N)$, and we find that the product expansion has the form $p^{-1}(1 - pq^{-1/N})$ times a product for which m and n are strictly positive. Thus, the coefficients vanish, and there are no roots of norm zero.

□

When g is non-Fricke, the last condition is somewhat more complicated to verify due to the presence of norm zero simple roots. Since the non-Fricke case is not needed in this paper, we postpone it to future work.

The following lemma will be useful in the next section, when we consider concrete quantitative questions. It is well-known to experts, but we couldn't find a statement in the literature that was both precise and concise.

Lemma 3.4.4. *Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be Borcherds-Kac-Moody Lie algebras with finite dimensional Cartan subalgebras $\mathfrak{h}_1, \mathfrak{h}_2$. Given an inner-product-preserving isomorphism $f : \mathfrak{h}_1 \to \mathfrak{h}_2$, there exists an extension to a Lie algebra isomorphism $\mathfrak{g}_1 \to \mathfrak{g}_2$ if and only if the root multiplicities are identical under the isometry f.*

Proof. It is clear that an isomorphism of Lie algebras that restricts to an isometry of Cartan subalgebras preserves root multiplicities.

Now, assume $\mathfrak{g}_1$ and $\mathfrak{g}_2$ have identical root multiplicities under the isometry $\mathfrak{h}_1 \to \mathfrak{h}_2$. It is well-known that the Weyl-Kac-Borcherds denominator formula may be used to uniquely determine the simple roots of a Borcherds-Kac-Moody algebra from the root multiplicities. Indeed, the set of values of $\alpha(h)$ as α ranges over positive roots is a discrete subset of the positive reals (in particular, well-ordered), and this allows us to recursively match the truncated evaluations of both sides of the Weyl-Kac-Borcherds denominator formula by adding simple roots as necessary. Indeed, this is essentially the same reconstruction process as the one sees in the proof of Lemma 3.4.2. Thus, under a choice of matching regular elements, $\mathfrak{g}_1$ and $\mathfrak{g}_2$ have the same simple roots, i.e., the simple roots that are identified under the isometry $\mathfrak{h}_1 \to \mathfrak{h}_2$ have equal multiplicity. Abstractly, this implies that both $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic to the Borcherds-Kac-Moody Lie algebra given by generators and relations from the same generalized Cartan matrix. More concretely, the simple root spaces are isomorphic vector spaces, and any choice of isomorphism between simple root spaces that are identified through the isometry $\mathfrak{h}_1 \to \mathfrak{h}_2$ extends uniquely to a Lie algebra isomorphism. □

Remark 3.4.5. Theorem 7.2 of [Borcherds-1992] is essentially an application of the previous lemma to identify the Monster Lie algebra $M = \mathfrak{m}_1$ defined by applying the no-ghost theorem to $V^\natural$ with the abstract Lie algebra $N = L_1$ defined by a specification of simple roots.

4. Comparison theorems

4.1. **Lie algebra comparison.** In [Carnahan-2012] section 4, we defined for each $g \in \mathbb{M}$ a Borcherds-Kac-Moody Lie algebra L_g (written W_g in Proposition 4.4 of that paper), whose Weyl-Kac-Borcherds denominator identity is an infinite product expansion of $T_g(\sigma) - T_g(-1/\tau)$. We now show that the quantization functor yields an isomorphic Lie algebra.

Proposition 4.1.1. *Let $g \in \mathbb{M}$ be a Fricke element. Then there exists a homogeneous Lie algebra isomorphism $\mathfrak{m}_g \to L_g$.*

Proof. By Proposition 3.4.3, $\mathfrak{m}_g$ is a Borcherds-Kac-Moody Lie algebra whose Cartan subalgebra has a real Lorentzian structure. By Theorem 4.2 of [Carnahan-2012], L_g is a Borcherds-Kac-Moody Lie algebra whose Cartan subalgebra has a real Lorentzian structure. By Lemma 3.4.4, if we are given two Borcherds-Kac-Moody Lie algebras and an inner-product-preserving isomorphism between their Cartan subalgebras, this isomorphism can be extended to an isomorphism of the full Lie algebras if and only if the root multiplicities are equal.

If $T_g(\tau)$ has level N, then the root lattices of both Lie algebras are identified with $II_{1,1}(-1/N)$, so it suffices to compare dimensions of the homogeneous spaces of nonzero degree. The dimensions are equal by the last claim in Proposition 3.2.6, which gives a $II_{1,1}(-1/N)$-graded vector space isomorphism $\mathfrak{m}_g \to L_g$. Thus, we have equality of dimensions of root spaces, and hence an isomorphism of Lie algebras. □

Corollary 4.1.2. *For any Fricke $g \in \mathbb{M}$, hypothesis C_g from section 4.4 of [Carnahan-2012] is satisfied, i.e., there is an action of $\widetilde{C_{\mathbb{M}}(g)}$ on L_g by Lie algebra automorphisms, such that the action on the $(a, b/N)$ root space coincides with the action on the subspace of $V^\natural(g^a)$ on which $\tilde{g}$ acts by $e(b/N)$ and L_0 acts by $1 + ab/N$.*

Proof. If we fix a Lie algebra isomorphism $\mathfrak{m}_g \to L_g$, then the action of $\widetilde{C_{\mathbb{M}}(g)}$ on L_g is given by *transport de structure*. The isomorphism between $\widetilde{C_{\mathbb{M}}(g)}$-module structures for the root spaces and the subspaces of twisted modules then follows from the corresponding fact for $\mathfrak{m}_g$, which holds by the fourth listed claim of Proposition 3.2.6. □

4.2. **Main theorem.** The following theorem is the target of this series of papers.

Theorem 4.2.1. *If g is a Fricke element, then for any $h \in C_{\mathbb{M}}(g)$, there exists a discrete group $\Gamma_{g,h} \subset PSL_2(\mathbb{R})$ containing a congruence subgroup $\Gamma(K)$ for some $K > 0$ (hence commensurable with $SL_2(\mathbb{Z})$), such that for any lift $\tilde{h}$ of h to a linear transformation on $V(g)$, the power series $Tr(\tilde{h}q^{L_0-1}|V^\natural(g))$ is the q-expansion of a modular function on $\mathcal{H}$ that is invariant under $\Gamma_{g,h}$, and generates the function field of the quotient $\mathcal{H}/\Gamma_{g,h}$.*

Proof. By Corollary 4.1.2, hypothesis C_g holds, and in particular the positive subalgebra E_g of L_g is Fricke-compatible for the data $(\widetilde{C_{\mathbb{M}}(g)}, \tilde{g}, \{V_k^{i,j} = V^\natural(g^i)^{\tilde{g}=e(j/N), L_0-1=k}\})$ in the sense of [Carnahan-2012] section 4.2. That is:

1. E_g is graded by $\mathbb{Z}_{>0} \times \frac{1}{N}\mathbb{Z}$, and decomposes as a sum $\bigoplus_{n>0, m\in\frac{1}{N}\mathbb{Z}} E_{i,j}p^iq^j$ of finite dimensional homogeneous components, where p and q denote degree shifts by $(0,1)$ and $(1,0)$.
2. E_g admits an action of $\widetilde{C_{\mathbb{M}}(g)}$ by homogeneous Lie algebra automorphisms, such that $E_{i,j} \cong V_{1+ij}^{i,j}$ as $\widetilde{C_{\mathbb{M}}(g)}$-modules.
3. The homology of E_g is given by
    - $H_0(E) = \mathbb{C}$
    - $H_1(E) = \bigoplus_{n\in\frac{1}{N}\mathbb{Z}} V_{1+n}^{1,n}pq^n$
    - $H_2(E) = p\bigoplus_{m=1}^{\infty} V_{1-1/N}^{1,-1/N} \otimes V_{1+m/N}^{m,1/N}p^m$
    - $H_i(E) = 0$ for $i > 2$.
4. $E_{1,-1/N} \cong V_{1-1/N}^{1,-1/N}$ is one dimensional.

By Proposition 6.3 in [Carnahan-2010], if E_g is Fricke-compatible for $(\widetilde{C_{\mathbb{M}}(g)}, \tilde{g}, \{V_k^{i,j}\})$, then for any $\tilde{h} \in \widetilde{C_{\mathbb{M}}(g)}$, the power series $Tr(\tilde{h}q^{L_0-1}|V^\natural(g))$ is the q-expansion of a Hauptmodul whose invariance group contains a principal congruence subgroup. □

By combining our main theorem with previously known results, we obtain a positive resolution of Norton's conjecture.

Theorem 4.2.2. *(Generalized Moonshine for twisted modules of $V^\natural$) Any rule that assigns*

- *to each $g \in \mathbb{M}$ the irreducible g-twisted $V^\natural$-module $V^\natural(g)$, with its canonical projective $C_{\mathbb{M}}(g)$-action, and*
- *to each commuting pair (g,h) the function $Z(g,h;\tau) = \mathrm{Tr}(\tilde{h}q^{L_0-1}|V^\natural(g))$ for some lift $\tilde{h}$ of h to a linear transformation on $V^\natural(g)$*

satisfies the following conditions:

1. *The formal series defining $Z(g,h;\tau)$ is the q-expansion of a holomorphic function on the upper half-plane $\mathfrak{H}$.*
2. *The function $(g,h) \mapsto Z(g,h;\tau)$ is invariant under simultaneous conjugation on the pair (g,h) in $\mathbb{M}$, up to rescaling.*
3. *$Z(g,h;\tau)$ is either a constant or a Hauptmodul.*
4. *For any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ and any commuting pair (g,h) in $\mathbb{M}$, $Z(g,h,\frac{a\tau+b}{c\tau+d})$ is proportional to $Z(g^ah^c, g^bh^d, \tau)$.*
5. *$Z(g,h;\tau)$ is proportional to $J(\tau)$ if and only if $g = h = 1$.*

Proof. The first claim is well-known: As a statement about normal convergence of power series, it follows from Theorem 8.1 of [Dong-Li-Mason-2000].

The second claim is also well-known, and follows by general nonsense from the Dong-Li-Mason proof of existence and uniqueness for irreducible twisted $V^\natural$-modules: For any $k \in \mathbb{M}$, k-conjugation on twisted modules induces an equivalence from the category of g-twisted $V^\natural$-modules to the category of kgk^{-1}-twisted $V^\natural$-modules, taking the projective $C_{\mathbb{M}}(g)$-action on $V^\natural(g)$ to the projective $C_{\mathbb{M}}(kgk^{-1})$-action on $V^\natural(kgk^{-1})$ by conjugation. This preserves characters up to rescaling.

The third claim is relatively straightforward in light of Theorem 4.2.1 combined with the $SL_2(\mathbb{Z})$ claim. If $g^a h^c$ is Fricke for some a, c satisfying $(a, c) = 1$, then $Z(g^a h^c, g^b h^d, \tau)$ is a Hauptmodul for all b, d such that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$. The $SL_2(\mathbb{Z})$ claim then implies $Z(g, h, \tau)$ is proportional to a Möbius transform of $Z(g^a h^c, g^b h^d, \tau)$, hence is a Hauptmodul. On the other hand, if all $g^a h^c$ for $(a, c) = 1$ are non-Fricke, then the twisted modules $V^\natural(g^a h^c)$ have L_0-spectrum bounded below by 1, so $Z(g^a h^c, g^b h^d, \tau)$ is regular at infinity for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$. Then the $SL_2(\mathbb{Z})$ claim implies $Z(g, h, \tau)$ is regular at all cusps, so the maximum modulus principle implies it is a constant function.

As we remarked in the introduction, the $SL_2(\mathbb{Z})$ claim follows immediately from combining Theorem 10.1 in [Dong-Li-Mason-2000] (which reduces the question to a g-rationality condition) with Corollary 5.26 of [Carnahan-Miyamoto-2016] (which asserts g-rationality).

The fifth claim also follows from [Dong-Li-Mason-2000]. In particular, the positivity of the L_0-spectrum of $V^\natural(g)$ for $g \neq 1$ (as given in Theorem 2.2.9) eliminates all nontrivial twistings. For the trivial twisting, the fact that the McKay-Thompson series $T_h(\tau) = Z(1, h, \tau)$ is equal to J if and only if $h = 1$ follows from the faithfulness of the monster action on $V^\natural$ as shown in [Frenkel, et al, 1988]. ☐

REFERENCES

[Borcherds-Conway-Queen-Sloane-1984]	R. E. Borcherds, J. H. Conway, L. Queen, N. J. A. Sloane, *A Monster Lie algebra?* Adv. in Math. **53** (1984) 75–79.
[Borcherds-1986]	R. Borcherds, *Vertex algebras, Kac-Moody algebras, and the Monster* Proc. Nat. Acad. Sci. USA, **83** no. 10 (1986) 3068–3071.
[Borcherds-1988]	R. Borcherds, *Generalized Kac-Moody algebras* J. Algebra, **115** no. 2 (1988) 501–512.
[Borcherds-1990]	R. Borcherds, *The Monster Lie algebra* Adv. Math. **83** no. 1 (1990) 30–47.
[Borcherds-1992]	R. Borcherds, *Monstrous moonshine and monstrous Lie superalgebras* Invent. Math. **109** (1992), 405–444.
[Borcherds-1995]	*A characterization of generalized Kac-Moody algebras* J. Algebra, **174** no. 3 (1995) 1073-1079.
[Brower-1972]	R. Brower, *Spectrum generating algebra and the no-ghost theorem for the dual model* Phys. Rev. **D6** (1972) 1655–1662.
[Carnahan-2007]	S. Carnahan, *Monstrous Lie algebras and Generalized Moonshine* Ph.D. Dissertation, U.C. Berkeley, May 2007.
[Carnahan-2010]	S. Carnahan, *Generalized Moonshine I: Genus zero functions* Algebra and Number Theory **4** no.6 (2010) 649–679. Available as `http://arxiv.org/abs/0812.3440`
[Carnahan-2012]	S. Carnahan, *Generalized Moonshine II: Borcherds products* Duke Math J. **161** no. 5 (2012) 893–950. Available as `http://arxiv.org/abs/0908.4223`
[Carnahan≥2016]	S. Carnahan, *Generalized Moonshine III: Equivariant intertwining operators* In preparation.
[Carnahan-Miyamoto-2016]	S. Carnahan, M. Miyamoto, *Regularity of fixed-point vertex operator subalgebras* ArXiv preprint `http://arxiv.org/abs/1603.05645`
[Conway-Norton-1979]	J. Conway, S. Norton, *Monstrous Moonshine* Bull. Lond. Math. Soc. **11** (1979) 308–339.
[Cummins-Gannon-1997]	C. Cummins, T. Gannon, *Modular equations and the genus-zero property of modular functions* Invent. Math. **129** (1997) 413–443.
[D'Hoker-1997]	E. D'Hoker, *String theory* Quantum Fields and Strings: a course for mathematicians, volume 2, ed. Deligne, *et al*, AMS, Providence, RI (1999) 807–1012.
[Dijkgraaf-Pasquier-Roche-1990]	R.Dijkgraaf, V. Pasquier, P. Roche, *Quasi Hopf algebras, group cohomology and orbifold models* Nucl. Phys. B (Proc. Suppl.) **18** B (1990) 60–72.

[Dixon-Ginsparg-Harvey-1988] L. Dixon, P. Ginsparg, J. Harvey, *Beauty and the Beast: Superconformal Symmetry in a Monster Module* Commun. Math. Phys. **119** (1988) 221–241.

[Dong-1994] C. Dong, *Representations of the moonshine module vertex operator algebra* Contemporary Math. **175** (1994) 27–36.

[Dong-Lepowsky-1993] C. Dong, J. Lepowsky, *Generalized vertex algebras and relative vertex operators* Progress in Mathematics **112** Birkhäuser Boston, Inc., Boston, MA, (1993).

[Dong-Li-Mason-2000] C. Dong, H. Li, G. Mason, *Modular invariance of trace functions in orbifold theory* Comm. Math. Phys. **214** (2000) no. 1, 1–56. Available as `http://arxiv.org/abs/q-alg/9703016`.

[Duncan-Frenkel-2011] J. F. R. Duncan, I. B. Frenkel, *Rademacher sums, moonshine and gravity* Commun. Number Theory Phys. **5** (2011), no. 4, 1–128. Available as `http://arxiv.org/abs/0907.4529`

[Eilenberg-Mac Lane-1954] S. Eilenberg, S. Mac Lane, *On the groups* $H(\Pi, n)$*, II, Methods of computation* Ann. Math. (2) **60** no. 1 (1954), 49–139.

[van Ekeren, et al, 2015] J. van Ekeren, S. Möller, N. Scheithauer, *Construction and classification of vertex operator algebras* ArXiv preprint `http://arxiv.org/abs/1507.08142`

[Feigin-1984] B. L. Feigin, *The semi-infinite homology of Kac-Moody and Virasoro Lie algebras* Uspekhi Mat. Nauk, **39**:2 (236) (1984), 195–196. Available at `http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=rm&paperid=2354`

[Feigin-Fuchs-1984] B. L. Feigin, D. B. Fuchs, *Verma modules over the Virasoro algebra* Topology, Proc. Internat. Topol. Conf. Leningrad 1982, L. D. Faddeev (ed.) A. A. Mal'tsev (ed.), Lect. notes in math. 1060, Springer (1984) 230–245.

[Frenkel-1985] I. Frenkel, *Representations of Kac-Moody algebras and dual resonance models* Applications of group theory in theoretical physics, Lect. Appl. Math. **21** A.M.S. (1985) 325–353.

[Frenkel-Ben-Zvi-2004] E. Frenkel, D. Ben-Zvi, *Vertex algebras and Algebraic Curves* Mathematical Surveys and Monographs **88** American Mathematical Society, Providence, RI, (2004).

[Frenkel-Garland-Zuckerman-1986] I. Frenkel, H. Garland, G. Zuckerman, *Semi-infinite cohomology and string theory* Proc. Nat. Acad. Sci. USA, **83** (1986) 8442–8446.

[Frenkel, et al, 1993] I. Frenkel, Y. Huang, J. Lepowsky, *On axiomatic approaches to vertex operator algebras and modules* Mem. Amer. Math. Soc. **104** (1993) no. 494.

[Frenkel, et al, 1988] I. Frenkel, J. Lepowsky, A. Meurman, *Vertex operator algebras and the Monster* Pure and Applied Mathematics **134** Academic Press, Inc., Boston, MA, (1988).

[Gaberdiel, et al, 2012] M. Gaberdiel, D. Persson, H.Ronellenfitsch, R.Volpato, *Generalized Mathieu moonshine* Commun.Num.Theor.Phys. **7** (2013) 145–223. Available at `http://arxiv.org/abs/1211.7074`

[Gaberdiel-Persson-Volpato-2013] M. R. Gaberdiel, D. Persson, R. Volpato, *Generalised Moonshine and Holomorphic Orbifolds* Proceedings, String-Math 2012, Bonn, Germany, July 16-21, 2012, (2013) 73–86. Available at `http://arxiv.org/abs/1302.5425`

[Ganter-2009] N. Ganter, *Hecke operators in equivariant elliptic cohomology and generalized moonshine* Groups and symmetries: from the Neolithic Scots to John McKay, CRM Proc. Lecture Notes **47** Amer. Math. Soc., Providence, RI (2009) 173–209. Available at `http://arxiv.org/abs/0706.2898`

[Goddard-Thorn-1972] P. Goddard, C. Thorn, *Compatibility of the dual pomeron with unitarity and the absence of ghosts in the dual resonance model* Physics Letters **40B** no. 2, (1972) 235–238.

[Green-Schwarz-Witten-1987] M. Green, J. Schwarz, E. Witten, *Superstring theory* Cambridge University Press, Cambridge, UK (1987).

[Höhn-2003b] G. Höhn, *Generalized moonshine for the Baby Monster* Preprint (2003)

[Huang-2008] Y.-Z. Huang, *Vertex operator algebras and the Verlinde conjecture* Commun. Contemp. Math., **10** no. 1 (2008) 103–154. Available at `http://arxiv.org/abs/math/0406291`

[Ivanov-Tuite-2002] R. I. Ivanov, M. P. Tuite, *Rational generalized moonshine and abelian orbifolds* Nucl. Phys. B **635** (2002) 435–472. Available at `http://arxiv.org/abs/math/0106027`

[Jurisich-1996] E. Jurisich, *An exposition of generalized KacMoody algebras* Lie algebras and their representations (Seoul, 1995), Contemporary Mathematics, **194** American Mathematical Society, Providence, RI (1996) 121–159.

[Jurisich-1998] E. Jurisich, *Generalized Kac-Moody Lie algebras, free Lie algebras and the structure of the Monster Lie algebra* J. Pure and Appl. Alg. **126** (1998) 233–266. Available at `http://arxiv.org/abs/1311.3258`

[Jurisich-2004] E. Jurisich, A resolution for standard modules of Borcherds Lie algebras *J. Pure Appl. Alg.* **192** (2004) 149–158.

[Kac-1997] V. Kac, *Vertex Algebras for Beginners* University Lecture Series **10**, Amer. Math, Soc., Providence, RI (1997)

[Kac-Raina-1987] V. Kac, A. Raina, *Bombay Lectures on highest weight representations of infinite dimensional Lie algebras* Advanced Series in Mathematical Physics **2**, World Scientific, Singapore, (1987).

[Kato-Ogawa-1983] M. Kato, K. Ogawa, *Covariant quantization of string based on BRS invariance* Nucl. Phys. B212 (1983) 443–460.

[Li-1994] H. Li *Symmetric invariant bilinear forms on vertex operator algebras* J. Pure Appl. Algebra, **96** no.3 (1994) 279–297.

[Lian-Zuckerman-1991] B. Lian, G. Zuckerman, *BRST Cohomology and highest weight vectors I* Comm. Math. Phys. **135** (1991) 547–580.

[Lian-Zuckerman-1993] B. Lian, G. Zuckerman, *New perspectives on the BRST-algebraic structure of string theory* Comm. Math. Phys. **154** (1993) 613–646. Available as `http://arxiv.org/abs/hep-th/9211072`

[Lian-Zuckerman-1995] B. Lian, G. Zuckerman, *Moonshine Cohomology* RIMS Proceedings on Vertex Operator Algebras and the Monster (1995). Available as `http://arxiv.org/abs/q-alg/9501025`

[Lovelace-1971] C. Lovelace *Pomeron form factors and dual regge cuts* Phys. Lett. **34B** (1971) 500–506.

[MacLane-1952] S. Mac Lane, *Cohomology theory of abelian groups* Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 2, Amer. Math. Soc., Providence, R.I., (1952) 8–14.

[Miyamoto-2013] *C_2-cofiniteness of cyclic-orbifold models* ArXiv preprint: `http://arxiv.org/abs/1306.5031`

[Norton-1987] S. Norton, *Generalized moonshine* Proc. Sympos. Pure Math. **47** Part 1, The Arcata Conference on Representations of Finite Groups (Arcata, Calif., 1986), 209–210, Amer. Math. Soc., Providence, RI (1987).

[Norton-2001] S. Norton, *From Moonshine to the Monster* Proceedings on Moonshine and related topics (Montréal, QC, 1999), 163–171, CRM Proc. Lecture Notes **30**, Amer. Math. Soc., Providence, RI (2001).

[Polchinski-1998] J. Polchinski, *String Theory* Cambridge University Press, Cambridge, UK (1998).

[Queen-1981] L. Queen, *Modular Functions arising from some finite groups* Mathematics of Computation **37** (1981) No. 156, 547–580.

[Scheithauer-1998] N. Scheithauer, *Vertex Algebras, Lie Algebras, and Superstrings* J. Algebra **200** (1998) 363–403. Available at `http://arxiv.org/abs/hep-th/9802058`

[Scheithauer-2000] N. Scheithauer, *The Fake Monster Superalgebra* Adv. Math. **151** (2000) 226–269. Available at `http://arxiv.org/abs/math/9905113`

[Zhu-1996] Y. Zhu, *Modular invariance of characters of vertex operator algebras* J. Amer, Math. Soc. **9** (1996) 237–302.

[Zuckerman-1989] G. Zuckerman, *Modular forms, strings, and ghosts* Proceedings of Symposia in Pure Mathematics, **49** I (1989) 273–284.

Division of Mathematics, University of Tsukuba
E-mail address: `carnahan@math.tsukuba.ac.jp`

arXiv:hep-th/9510182v2 9 Jan 1996

hep-th/9510182
EFI-95-64
YCTP-P16-95

Algebras, BPS States, and Strings

Jeffrey A. Harvey

Enrico Fermi Institute, University of Chicago
5640 Ellis Avenue, Chicago, IL 60637
harvey@poincare.uchicago.edu

Gregory Moore

Department of Physics, Yale University
New Haven, CT 06511
moore@castalia.physics.yale.edu

Abstract

We clarify the role played by BPS states in the calculation of threshold corrections of D=4, N=2 heterotic string compactifications. We evaluate these corrections for some classes of compactifications and show that they are sums of logarithmic functions over the positive roots of generalized Kac-Moody algebras. Moreover, a certain limit of the formulae suggests a reformulation of heterotic string in terms of a gauge theory based on hyperbolic algebras such as E_{10}. We define a generalized Kac-Moody Lie superalgebra associated to the BPS states. Finally we discuss the relation of our results with string duality.

October 24, 1995, Revised Jan. 9, 1996

1. Introduction

It has become clear recently that theories with extended supersymmetry have a rich dynamical structure which is nonetheless amenable to exact analysis using ideas of electromagnetic duality [1,2,3,4,5,6,7,8]. One important feature of these theories is the existence of BPS states. These states play a crucial role in the dynamics of the theory and their structure at strong coupling can in many cases be determined by a semi-classical analysis. For example, in the analysis of [5,6] certain BPS states become massless at special points in the quantum moduli space of vacua and dominate the low-energy dynamics.

In string theory there is increasing evidence that a related although undoubtedly richer structure is present, particularly in $N = 2$ theories exhibiting duality [9,10]. String theories with extended supersymmetry resulting from toroidal compactification have in fact an infinite spectrum of BPS states [11,12]. These BPS states have played a central role in much of the recent work on duality in string theory. In Type II string theory the presence of an infinite tower of non-perturbative BPS states is crucial in understanding the strong coupling behavior and duality symmetries [13,14]. These BPS states carrying Ramond-Ramond charge are also essential to the resolution of the conifold singularity in Type II string theory [15,16] and play an important role in understanding the relation between Type I and Type II string theory [17]. In the heterotic string BPS states played a central role in the original understanding of S-duality [18]. It seems fair to say that the foundations of string theory are shifting and that the structure of BPS states provides one of the most useful clues as to what type of theory unifies these disparate phenomena.

Another theme which runs through much of string theory is the search for the symmetries which underlie the structure of string theory. One symmetry structure which has been investigated in this regard is that of hyperbolic Kac-Moody and generalized Kac-Moody algebras [19,20,21,22,23,24,25,26] . Generalized Kac-Moody (GKM) algebras occur very naturally as unbroken symmetry groups in certain string ground states [23] and -contrary to what one expects in spontaneously broken gauge theory - even when these gauge symmetries are broken they nevertheless put strong constraints on the S-matrix [24]. The idea of a T-duality invariant string algebra based on a Lorentzian lattice as a gauge algebra for $N = 4$ heterotic string compactifications was discussed in [22] and is summarized in [27]. There are also close connections between these algebras and the structure of string field theory [28].

In this paper we will provide evidence for a connection between BPS states in string theory with $N = 2$ spacetime supersymmetry and GKM algebras. We will show that threshold corrections in $N=2$ heterotic string compactifications are in fact determined purely in terms of the spectrum of BPS states. What is more surprising is that these corrections are closely related to product formulae that have been studied recently by Borcherds, Gritsenko and Nikulin [29,30,31,32,33] in connection with generalized Kac-Moody algebras. We will show how this connection arises and construct a GKM Lie superalgebra in terms of vertex operators associated to BPS states.

Our results can be viewed as a generalization to string theory of some of the structures previously encountered in duality and $N=2$ Yang-Mills theory. For example, our results suggest that the BPS states in string theory should be regarded as "gauge bosons " of the GKM algebra. We find a direct generalization of the one-loop formula for the prepotential in $N=2$ Yang-Mills theory [7,8],

$$\mathcal{F} = \frac{i}{4\pi} \sum_{\vec{\alpha}>0} (\alpha \cdot A)^2 \log \frac{(\alpha \cdot A)^2}{\Lambda^2} \tag{1.1}$$

where A determines the components of the Higgs expectation value in the Cartan subalgebra $\phi = \sum A_i H_i$ and the sum in (1.1) runs over the positive roots of the Lie algebra of the gauge group G. In a certain limit we will find a similar formula in string theory where the sum over the positive roots of the Lie algebra of G is replaced by a sum over the positive roots of a GKM algebra.

The simplest example of a product formula which occurs in threshold corrections is Borcherds' remarkable product formula for the modular j function

$$j(p) - j(q) = p^{-1} \prod_{n>0, m\in Z} (1-p^n q^m)^{c(mn)} \tag{1.2}$$

where $j(q) - 744 = \sum_{n=-1} c(n) q^n$. The proof can be found in [34]and [35]. The infinite product converges for $|p|$, $|q| < e^{-2\pi}$ and is defined elsewhere by analytic continuation. An expansion of both sides for small p, q quickly reveals that it requires an infinite number of identities among the coefficients, the first of which is $c(4) = c(3) + c(1)(c(1)-1)/2$. In [36] these identities appear as "replication" formulae involving the dimensions of irreducible representations of the monster group. In [35] the product formula (1.2) is interpreted in terms of the denominator formula for the monster Lie algebra which is an example of a GKM algebra. The quantity $j(T) - j(U)$ has appeared in the recent literature in connection with threshold corrections to gauge couplings with T, U the moduli on a T^2

in the string compactification [37,38,39]. [1] In [37,38,39] this term was found by indirect methods. We will show here how it and various generalizations can be computed directly in terms of products which generalize (1.2). Indeed, using the integrals of appendix A it is straightforward to give an independent proof of (1.2).

There is a substantial literature on threshold corrections in $N = 1$ and $N = 2$ heterotic string theory [40,41,42,43,44,45,46,37,39,38,47,48]. We will make particular use of some techniques of [42] and of the result in [43][48] relating threshold corrections to the new supersymmetric index [49]. We also rely heavily on the results of [38,39] relating explicit one-loop string calculations to the quantities appearing in $N = 2$ supergravity effective Lagrangians. A connection between BPS states and threshold corrections has been suggested, directly and indirectly, in various forms, in [50,51,47,52].

The outline of this paper is as follows. In the second section we briefly review the relevant properties of $N = 2$ heterotic string compactifications and associated moduli spaces. In the third section we discuss the nature of perturbative BPS states in these compactifications. We show that threshold corrections are determined purely by the spectrum of BPS states and relate these corrections to the elliptic genus of the internal $c = 6$, $N = 4$ superconformal field theory governing the compactification. In the fourth section we determine the dependence of these corrections on either the (T, U) Narain moduli of T^2 or these plus the E_8 Wilson line moduli and show that they are naturally given by infinite sums of logarithmic functions. The fifth section contains a discussion of the relation of our work to a theorem of Borcherds. We discuss the automorphic structure of the products which result from our analysis and show how Borcherds results follow from an analysis of certain modular integrals. In section six we discuss the quantum monodromy of the prepotential. The seventh section contains a brief discussion of GKM algebras and their associated root lattices and compares the results of sections four and five to the denominator formula for these algebras. For the previous two choices of moduli we show that the product formulae we obtain are given by a product over positive roots of the monster Lie algebra and the hyperbolic Kac-Moody Lie algebra E_{10} respectively. We discuss two limiting cases of the prepotential in section eight, one of which suggests that the GKM algebra should be considered a gauge algebra in string theory. In the ninth section we construct a generalized Kac-Moody algebra which is in some sense associated to the BPS states. The tenth section has some preliminary remarks on the application of our results to $N = 2$ string duality. In the final section we conclude and offer some speculations on possible extensions of this work. Certain modular integrals needed in the evaluation of threshold corrections are computed in an appendix.

---

[1] We use both notations $j(q)$ and $j(\tau)$, depending on context.

2. Review of $D = 4, N = 2$ heterotic string compactifications

2.1. Chiral Algebra

In this paper we will be considering compactifications of the heterotic string to four spacetime dimensions with $N = 2$ spacetime supersymmetry. There is a $c = 9$, $N = 1$ internal superconformal algebra (SCA) associated to any such compactification. For such theories the spacetime supersymmetry implies that this internal SCA splits into a $c = 6$ piece with $N = 4$ superconformal symmetry and a $c = 3$ piece with $N = 2$ superconformal symmetry [53]. The $c = 3$, $N = 2$ theory is constructed from two free dimension 1/2 superfields. We will indicate this decomposition of the right-moving superconformal algebra as

$$\tilde{\mathcal{A}}^{N=2}_{\tilde{c}=3} \oplus \tilde{\mathcal{A}}^{N=4}_{\tilde{c}=6} \subset \tilde{\mathcal{A}} \quad , \tag{2.1}$$

where $\tilde{\mathcal{A}}$ is the rightmoving chiral algebra. The left-moving internal conformal field theory has $c = 22$ but is otherwise unconstrained except by modular invariance.

The $c = 3$ theory has a $U(1)$ current which we denote by $J^{(1)}$. The $c = 6$ theory is in general non-trivial and to be compatible with $N = 4$ superconformal symmetry must have a level one $SU(2)$ Kac-Moody algebra. Representations of the $c = 6$ theory can thus be labelled by the conformal weight and $SU(2)$ representation (h, I). We will also choose a $U(1)$ current $J^{(2)}$ from this $SU(2)$ algebra with the normalization $J^{(2)} = 2J^3$ where J^3 is the $SU(2)$ Cartan current. The total $U(1)$ current of the $c = 9$ theory is $J = J^{(1)} + J^{(2)}$.

2.2. Remarks on the moduli space

$N = 2$ heterotic compactifications typically have a large moduli space of vacua. The moduli decompose into hypermultiplets and vectormultiplets under the $N = 2$ spacetime supersymmetry. The $c = 6$, $N = 4$ SCA has two massless representations in the Neveu-Schwarz sector, $(h = 0, I = 0)$ and $(h = 1/2, I = 1/2)$ [54]. These are associated to vectormultiplet and hypermultiplet moduli respectively. The factorization (2.1) is reflected in spacetime through the constraints of $N = 2$ supergravity which require a local factorization of the total moduli space of the form

$$SK(n) \times Q(m) \tag{2.2}$$

where $SK(n)$ is the vectormultiplet moduli space which must be a special Kahler manifold of real dimension $2n$ and $Q(m)$ is the hypermultiplet moduli space and is quaternionic of real dimension $4m$. In fact the moduli space is a complicated and singular space; its global structure has not been fully elucidated and should prove very interesting. The classical moduli space is an algebraic variety which probably has the following structure. $\mathcal{M}$ is a singular stratified space, that is, there is a disjoint union

$$\mathcal{M} = \amalg_{\int} \mathcal{M}_{\int} \tag{2.3}$$

with smooth strata $\mathcal{M}_{\int}$ which fit together in a singular fashion. The $\mathcal{M}_{\int}$ have the structure of a fibration of a quaternionic manifold over a special Kahler manifold. [2]

The classical vectormultiplet moduli space for $N = 2$ string compactifications is given by the special Kahler manifold [55]

$$\mathcal{M}_{\mathrm{vm}}^{s+2,2} \equiv \frac{SU(1,1)}{U(1)} \times \mathcal{N}^{s+2,2} \tag{2.4}$$

where the first factor is associated with the dilaton. Here we have introduced the Narain moduli space $\mathcal{N}^{s+2,2}$. This is a quotient of the generalized upper half plane

$$\mathcal{H}^{s+1,1} \equiv O(s+2,2;\mathbb{R})/[O(s+2)\times O(2)] \tag{2.5}$$

by the left-action of the arithmetic subgroup $O(s+2,2;\mathbb{Z})$:

$$\begin{aligned} \mathcal{N}^{s+2,2} &\equiv O(s+2,2;\mathbb{Z})\backslash\mathcal{H}^{s+1,1} \\ &= O(s+2,2;\mathbb{Z})\backslash O(s+2,2;\mathbb{R})/[O(s+2)\times O(2)] \end{aligned} \tag{2.6}$$

The special geometry of (2.6) is described in some detail below.

For $N = 2$ compactifications with internal space $K3 \times T^2$ and an embedding of the spin connection in $E_8 \times E_8$ the moduli space $\mathcal{M}_{\mathrm{vm}}^{s+2,2}$ always contains a subvariety $\cong \mathcal{N}^{2,2}$, which is the Narain moduli space for the complex moduli T, U on T^2:

$$\frac{O(2,2)}{O(2)\times O(2)} \simeq \left(\frac{SU(1,1)}{U(1)}\right)_T \otimes \left(\frac{SU(1,1)}{U(1)}\right)_U . \tag{2.7}$$

---

[2] In the case of global $N = 2$ SYM with matter these statements can probably be proven from the F and D flatness equations. For gauge group G the strata should be enumerated by the possible unbroken subgroups $G_1 \subset G$. The base should be $\underline{\mathbf{t}}(G_1) \otimes \mathbb{C}/W(G_1)$ while the fibers are hyperkahler quotients.

The arithmetic group $O(2,2;\mathbb{Z})$ in this case consists of the $T \leftrightarrow U$ exchange and transformations by $PSL(2,\mathbb{Z})_T \times PSL(2,\mathbb{Z})_U$.

In $N=2$ string theory the classical special geometry of $\mathcal{M}^{s+2,2}_{\rm vm}$ receives quantum corrections. The dependence of threshold corrections on T,U moduli has been well studied in the literature and has played an important role in recent tests of string-string duality [56,57,58]. In many compactifications based on $K3\times T^2$ one finds additional vectormultiplet moduli associated to Wilson lines for the unbroken gauge group on T^2. The dependence of threshold corrections on these Wilson line moduli has been studied in [37,38,59].

In this paper we will for the most part focus on two special cases with moduli space (2.4) for $s=0$ and $s=8$ corresponding to the T,U moduli on T^2 in the first case and in the second to T,U and the Wilson line moduli associated to the unbroken E_8 factor for $K3$ compactifications given by the standard embedding of the spin connection in the gauge group. For these two cases the moduli are just the Narain moduli associated to the even self-dual lattices $\Pi^{2,2}$ and $\Pi^{10,2}$. However many of our considerations are more general and we will only specialize to these two cases when necessary.

2.3. The classical special geometry of $\mathcal{N}^{s+2,2}$ *and of* $\mathcal{M}^{s+2,2}_{\rm vm}$.

The standard Kahler geometry on the space (2.4) is special geometry described, for example, in [38,60]. We now discuss a useful parametrization of the homogeneous space

$$\mathcal{H}^{s+1,1} = \frac{O(s+2,2)}{O(s+2)\times O(2)} \tag{2.8}$$

occurring in (2.4). Further details may be found in [61,62,38,60,37].

We can represent the homogeneous space $\mathcal{H}^{s+1,1}$ as a "tube domain" in a complexified Minkowski space as follows. Let $\langle\cdot,\cdot\rangle$ be a real quadratic form of signature $(+^{s+2},-^2)$. Consider:

$$\{u\in\mathbb{C}^{s+4} : \langle u,u\rangle = 0, \langle u,\bar u\rangle < 0\}/(u\sim\lambda u) \tag{2.9}$$

This is a homogeneous space for $O(s+2,2;\mathbb{R})$. We let $y\in\mathbb{R}^{s+1,1}\otimes\mathbb{C}$, where we have a real inner product $(,)$ on $\mathbb{R}^{s+1,1}$ of signature $(+^{s+1},-^1)$. We parametrize the solutions to (2.9) by

$$y\to u(y)\equiv\Bigl(y;1,-\tfrac{1}{2}(y,y)\Bigr) \tag{2.10}$$

where the inner products are related by

$$\langle v;x_1,x_2\rangle^2 \equiv (v,v)+2x_1x_2 \tag{2.11}$$

Moreover

$$\langle u(y), \overline{u(y)} \rangle = +2(\Im y, \Im y) \tag{2.12}$$

so we must have $(\Im y)^2 < 0$, i.e., a lightlike vector [3].

The lightcone has two components, take $\Im y \in C_+$, the forward light-cone, so we realize the moduli space as a generalized upper half plane:

$$O(s+2,2)/[O(s+2) \times O(2)] \cong \mathcal{H}^{s+1,1} \equiv \mathbb{R}^{s+1,1} + iC_+^{s+1,1} \tag{2.13}$$

Sometimes when we want to be more specific about the inner product in (2.10) we take

$$\begin{aligned} v &= (\vec{v}; v_+, v_-) \\ (v,v) &= \vec{v}^2 - 2v_+v_- \end{aligned} \tag{2.14}$$

where $\vec{v}^2$ is the standard Euclidean inner product.

One important special case is the submanifold $\vec{y} = 0$, isomorphic to the case $s = 0$. Then we have two complex numbers $\Im y_+ \Im y_- > 0$. The forward lightcone then gives

$$(y_+, y_-) \to (T, U) \in \mathcal{H} \times \mathcal{H} \tag{2.15}$$

in the product of two upper half planes.

The Kahler metric on $\mathcal{N}^{s+2,2}$ is

$$K = -\log[-\tfrac{1}{2}\langle u, \bar{u}\rangle] = -\log[-(\Im y^2)] = -\log[2\Im y_+ \Im y_- - (\Im \vec{y})^2]. \tag{2.16}$$

Note that the argument of the log is positive by the definition of the domain of y and that the resulting metric is one of constant negative curvature. Similarly, the classical prepotential and Kähler potential for the space (2.4) are $\mathcal{F} = -S(y,y)$ and $K = -\log[8\Im S] - \log[-(\Im y)^2]$, respectively.

For our purposes the space $\mathcal{N}^{s+2,2}$ arises as the moduli space for Narain compactifications based on the even self-dual Lorentzian lattice $\Pi^{8t+2,2}$ when $s = 8t$. [4] We write $\Pi^{8t+2,2} \cong \Pi^{8t,0} \oplus \Pi^{2,2}$ and write lattice vectors as

$$(\vec{b}; m_+, n_-; m_0, n_0) \tag{2.17}$$

with metric

$$(\vec{b}; m_+, n_-; m_0, n_0)^2 = \vec{b}^2 - 2m_+n_- + 2m_0n_0. \tag{2.18}$$

Here $\Pi^{8t,0}$ is an even self-dual Euclidean lattice. For $t = 0$ it is the zero lattice $\{\vec{b} = 0\}$. It is unique for $t = 1$, there are two for $t = 2$, twenty-four for $t = 3$, and so forth.

---

[3] We will use both $\Re z$, $\Im z$ and z_1, z_2 to indicate the real and imaginary parts of a complex number z in this paper

[4] We use the symbol $\Pi^{p,q}$ to denote a standard even unimodular Lorentzian lattice. Lattices isomorphic to it are denoted by $\Gamma^{p,q}(y)$, where $y \in \mathcal{N}^{p,q}$.

3. Supersymmetric Index and BPS states in $N = 2$ heterotic string compactifications

3.1. BPS States

Compactifications of the heterotic string with $N = 2$ spacetime supersymmetry have a spacetime supersymmetry algebra which includes a complex central charge $\mathcal{Z}$:

$$\{Q^i_\alpha, Q^j_\beta\} = \epsilon_{\alpha\beta}\epsilon^{ij}\mathcal{Z}. \tag{3.1}$$

The central charge $\mathcal{Z}$ is determined by the right-moving momenta p_R carried by the free superfields in the $c = 3$, $N = 2$ internal SCA. The Virasoro constraints imply that the mass of any state in this theory is given by (in the Neveu-Schwarz sector)

$$M^2 = (N_R - 1/2) + \tfrac{1}{2}p_R^2 + h_R \tag{3.2}$$

where N_R is the right-moving oscillator number coming from the uncompactified coordinates and the two free superfields of the $c = 3$ theory and h_R is the conformal weight in the $c = 6$ part of the internal SCA. As a result, bosonic BPS states which must satisfy $M^2 = \frac{1}{2}p_R^2$ arise as right-moving ground states with either $N_R = 1/2$ and $h_R = 0$ or $N_R = 0$ and $h_R = 1/2$.

For Narain compactifications with lattices $\Pi^{8t+2,2}$ and lattice vectors (2.17) we have

$$\begin{aligned}\tfrac{1}{2}(p_L^2 - p_R^2) &= \tfrac{1}{2}\vec{b}^2 - m_+n_- + m_0n_0\\ \tfrac{1}{2}p_R^2 &= \frac{1}{-2(\Im y)^2}\left|\vec{b}\cdot\vec{y} - m_+y_- - n_-y_+ + n_0 - \tfrac{1}{2}m_0y^2\right|^2.\end{aligned} \tag{3.3}$$

That is, the central charges are determined by the inner product of the lattice vector (2.17) and the coordinate $u(y)$.

In contrast to theories with $N = 4$ spacetime supersymmetry, the spectrum of BPS states in $N = 2$ theories can have a "chaotic" nature, that is BPS states can appear and disappear under infinitesimal perturbations of the moduli. One simple example which illustrates this is a symmetric orbifold limit of $K3$. We consider a Z_2 orbifold of a $\Pi^{4,4}$ lattice. In the untwisted sector of the orbifold we will have states labelled by momenta $(\tilde{p}_L, \tilde{p}_R) \in \Pi^{4,4}$. Such states can be BPS states only if $\tilde{p}_R = 0$. But the existence of such states varies discontinuously as we vary the Narain moduli associated to the $\Pi^{4,4}$ lattice. For example if we focus on one S^1 factor of radius R then states with $(\tilde{p}_L \neq 0, \tilde{p}_R = 0)$ exist only for rational values of the modulus $2R^2$.

This seems to be in contradiction with the usual wisdom that BPS states must behave smoothly under perturbations of the theory. The resolution of this puzzle is simply that these chaotic BPS states always appear in hypermultiplet, vectormultiplet pairs. As one moves away from the special points these BPS states pair into long representations of the $N = 2$ spacetime supersymmetry algebra and are no longer BPS saturated. This also makes it clear that threshold corrections cannot depend only on the number and charges of BPS states if one is to obtain smooth functions of the moduli. Rather, as we will see in the following section, the threshold corrections depend only on the difference between vectormultiplet and hypermultiplet BPS states and this difference is a smooth function of the moduli.

3.2. BPS states and the supersymmetric index

It was shown in [43] that threshold corrections in $N = 2$ heterotic string compactifications can be written in terms of the "new supersymmetric index" of [49]:

$$\mathrm{Tr_R J_0 e^{i\pi J_0} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24}} = \frac{1}{2\pi \mathrm{i}} \frac{\partial}{\partial \theta}\Big|_{\theta = \frac{1}{2}} \mathrm{Tr}_{\mathcal{H}^{\mathrm{int}}_{\mathrm{R}}} \mathrm{q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24} e^{2\pi i \theta J_0}} \tag{3.4}$$

where the trace is over the Ramond sector of the internal $(c, \tilde{c}) = (22, 9)$ conformal field theory and J_0 is the total $U(1)$ charge defined earlier. In the literature J_0 is often denoted by F.

We will now show that for $N = 2$ compactifications one can relate (3.4) to a sum over BPS saturated states. Using the decomposition (2.1)the Hilbert space of the internal superconformal field theory may be written as

$$\mathcal{H}^{\mathrm{int}} = \sum_{(p_R; h, I)} \mathcal{H}^{(22,0)}_{p_R; h, I} \otimes \tilde{\mathcal{H}}^{(0,3)}_{p_R} \otimes \tilde{\mathcal{H}}^{(0,6)}_{h, I} \tag{3.5}$$

where superscripts denote the Virasoro central charges $(c, \tilde{c})$, the second factor is a Fock space for the free N=2 superfield, and the third factor is a unitary irrep of the $N = 4$ SCA labelled by the conformal weight and $SU(2)$ representation (h, I). Each summand in (3.5) is a tensor product of representations of the $\tilde{\mathcal{A}}^{\tilde{c}=3}_{N=2}$ and the $\tilde{\mathcal{A}}^{\tilde{c}=6}_{N=4}$ algebras. Accordingly, the $U(1)$ current may be decomposed as $J = J^{(1)} + J^{(2)}$ and hence we may rewrite the trace on each summand in (3.5) as:

$$\begin{aligned}
&\mathrm{Tr}_{\mathrm{p_R \otimes (h,I)}} J_0 e^{i\pi J_0} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24} = \\
&\qquad \mathrm{Tr_{p_R} J_0^{(1)} e^{i\pi J_0^{(1)}} \bar{q}^{\tilde{L}_0 - \tilde{c}/24}} \left(\mathrm{Tr_{(h,I)} e^{i\pi J_0^{(2)}} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24}} \right) \\
&\qquad + \mathrm{Tr_{p_R} e^{i\pi J_0^{(1)}} \bar{q}^{\tilde{L}_0 - \tilde{c}/24}} \left(\mathrm{Tr_{(h,I)} J_0^{(2)} e^{i\pi J_0^{(2)}} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24}} \right)
\end{aligned} \tag{3.6}$$

Now, for an arbitrary $N=4$ representation (h,I) one has

$$\mathrm{Tr}_{(\mathrm{h},\mathrm{I})} \mathrm{J}_0^{(2)} \mathrm{e}^{\mathrm{i}\pi \mathrm{J}_0^{(2)}} \mathrm{q}^{\mathrm{L}_0 - \mathrm{c}/24} \bar{\mathrm{q}}^{\tilde{\mathrm{L}}_0 - \tilde{\mathrm{c}}/24} = 0. \tag{3.7}$$

To see this recall that the $U(1)$ current is related to the $SU(2)$ Cartan current by $J = 2J^3$ and hence has integral spectrum. Since in $SU(2)$ representations eigenvalues of J^3 come in opposite pairs (3.7) follows. Thus, only the first term on the right hand side of (3.6) survives.

We can further simplify (3.6) using N=4 representation theory. As we saw earlier BPS states correspond in the Neveu-Schwarz sector to operators in the $N=4$ theory with $h_R = 0, 1/2$. These are just the massless NS representations of the $N=4$ SCA with $(h=0, I=0)$ and $(h=1/2, I=1/2)$. The representation $(h=0, I=0)$ gives rise to the bosonic states in a BPS vectormultiplet. On the other hand the representation $(h=1/2, I=1/2)$ gives rise to the bosonic states of a BPS hypermultiplet[5].

Spectral flow in an $N=4$ theory maps these representations to Ramond representations according to

$$\begin{gathered} (h=0, I=0) \rightarrow (1/4, 1/2) \\ (h=1/2, I=1/2) \rightarrow (1/4, 0). \end{gathered} \tag{3.8}$$

These are the massless Ramond representations with Witten indices [54]:

$$\begin{gathered} \mathrm{Tr}_{\mathcal{H}^{\mathrm{N}=4}_{(\mathrm{h}=1/4,\mathrm{I}=0)}} (-1)^{2\mathrm{J}_0^3} = 1 \\ \mathrm{Tr}_{\mathcal{H}^{\mathrm{N}=4}_{(\mathrm{h}=1/4,\mathrm{I}=1/2)}} (-1)^{2\mathrm{J}_0^3} = -2 \end{gathered} \tag{3.9}$$

In fact, these are the only representations for which the Witten indices are nonzero [54]. It follows that the sum (3.4) reduces to a sum over BPS states.

The representation content of BPS multiplets with respect to the subalgebra $\tilde{\mathcal{A}}^{c=3}_{N=2} \oplus \tilde{\mathcal{A}}^{c=6}_{N=4}$ may be summarized as follows. We will denote a representation of this algebra by $(h,q) \otimes (h', I)$ where (h,q) give the conformal weight and $U(1)$ charge of the $c=3, N=2$ theory and (h', I) labels the representation of the $N=4$ algebra . With this understood vectormultiplets and hypermultiplets have the following content in the Neveu-Schwarz (NS) and Ramond (R) sectors:

---

[5] Our use of the terms "hypermultiplet" and "vectormultiplet" here is nonstandard. We define these terms based purely on the right-moving structure of the representation. For example, with this terminology the supergravity multiplet is counted as a "vectormultiplet."

$$\begin{array}{rcc} & \text{Vectormultiplets} & \text{Hypermultiplets} \\ \mathrm{NS}: & 2\times(0,0)\otimes(0,0)\oplus(1/2,\pm1)\otimes(0,0) & 2\times(0,0)\otimes(1/2,1/2) \\ \mathrm{R}: & (1/8,\pm1/2)\otimes(1/4,1/2) & 2\times(1/8,\pm1/2)\otimes(1/4,0) \end{array} \tag{3.10}$$

where all combinations of ± signs should be taken.

Combining (3.10) with (3.9) we see that each BPS vectormultiplet contributes

$$[\tfrac{1}{2}e^{i\pi/2}-\tfrac{1}{2}e^{-i\pi/2}](-2)=-2i \tag{3.11}$$

to $J_0 e^{i\pi J_0}$ while each BPS hypermultiplet contributes

$$2\times[\tfrac{1}{2}e^{i\pi/2}-\tfrac{1}{2}e^{-i\pi/2}](+1)=+2i \tag{3.12}$$

to $J_0 e^{i\pi J_0}$. Thus the BPS states contribute to (3.4) with vectormultiplets and hypermultiplets weighted with opposite signs, that is:

$$\begin{gathered} \frac{1}{\eta^2}\mathrm{Tr_R J_0 e^{i\pi J_0} q^{L_0-c/24}}\bar{q}^{\tilde{L}_0-\tilde{c}/24}= \\ -2i\left[\sum_{\text{BPS vectormultiplets}} q^{\Delta}\bar{q}^{\bar{\Delta}}-\sum_{\text{BPS hypermultiplets}} q^{\Delta}\bar{q}^{\bar{\Delta}}\right] \end{gathered} \tag{3.13}$$

Strictly speaking this equation is not completely correct because of the fact that non-physical states with $\Delta\neq\bar{\Delta}$ nonetheless contribute to modular integrals. With the understanding that the contribution of these non-physical BPS states should be included as well (3.13) is a correct equation.

Note that this is in accord with the expectation that there are no threshold corrections for $N=4$ spacetime supersymmetry and indeed $N=4$ BPS states split up into a $N=2$ hypermultiplet and a $N=2$ vectormultiplet. Another way to understand this result from the spacetime point of view is to notice that when representing the extended supertranslation algebra the massive *long* $N=2$ representations are the same as the *short* $N=4$ representations. But we know there are never any threshold corrections in $N=4$ theories, therefore we expect threshold corrections only from short $N=2$ representations, that is from BPS states.

3.3. *Elliptic genus*

From (3.6) and (3.7) we see that the new supersymmetric index (3.4) depends on $N = 4$ moduli only through

$$\mathrm{Tr}_{\mathrm{Ramond}} \mathrm{e}^{\mathrm{i}\pi \mathrm{J}_0^{(2)}} \mathrm{q}^{\mathrm{L}_0 - \mathrm{c}/24} \bar{\mathrm{q}}^{\tilde{\mathrm{L}}_0 - \tilde{\mathrm{c}}/24} \tag{3.14}$$

which, roughly speaking, is the elliptic genus [63,64,65] of the $N = 4$ conformal field theory. For a general class of backgrounds we can be more specific about the relation to the conventional elliptic genera. We assume that the gauge bundle on $T^2 \times K3$ has the structure $\pi_1^*(V_1) \oplus \pi_2^*(V_2)$, that is, in the fermionic formulation of the gauge algebra, the leftmoving fermions λ^I can be split into two disjoint sets coupling via their currents as

$$\lambda^I A_\mu^{IJ} \partial X^\mu \lambda^J + \lambda^I B_\mu^{IJ} \partial X^\mu \lambda^J \tag{3.15}$$

where A is a flat connection on T^2 for the bundle V_1, and B is an anti-self-dual instanton connection for the bundle V_2 on $K3$. We further assume that we can embed the connection B in an $SO(2n)$ subgroup of $SO(16) \subset E_8$. We can then use the bosonic formulation for one E_8 factor and the fermionic formulation for the E_8 factor containing the connection B and split the 16 fermions as $16 = 2n + (16 - 2n)$. We then couple the $2n$ fermions to the connection B via (3.15). In this case we have $(16 - 2n)$ free Majorana-Weyl fermions coupled to the flat connection A on T^2, and a $(c, \tilde{c}) = (4 + n, 6)$ heterotic sigma model on $K3$ with a gauge bundle

$$\begin{tikzcd} V_2 \arrow[d] \\ K3 \end{tikzcd} \tag{3.16}$$

satisfying the conditions

$$c_1(V_2) = 0, \qquad ch_2(V_2) = \tfrac{1}{2} c_1(V_2)^2 - c_2(V_2) = c_2(TK3) = +24. \tag{3.17}$$

We now define elliptic genera for the $(4 + n, 6)$ sigma-model as

$$\begin{aligned} \Phi^+(\hat{A}) &= Tr_{NS,R}(-1)^{F_R} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24} \\ \Phi^-(\hat{A}) &= Tr_{NS,R}(-1)^{F_L + F_R} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24} \\ \Phi(\Delta) &= Tr_{R,R}(-1)^{F_R} q^{L_0 - c/24} \bar{q}^{\tilde{L}_0 - \tilde{c}/24} \end{aligned} \tag{3.18}$$

where the subscript on the trace indicates the left-moving and right-moving boundary conditions respectively and F_L and $F_R (\equiv J_0^{(2)})$ are the left and right-moving fermion number.

The $\Phi^{\pm}$ are the elliptic generalization of the Dirac index [64] where the ± determines whether odd antisymmetric tensor representations of $SO(2n)$ are counted with a plus sign or minus sign. Φ^{+} is denoted by $H(q)$ in the first reference of [64]. The quantity $\Phi(\Delta)$ involves the elliptic generalization of the Dirac index coupled to the spinor bundles associated to V_2.

We can combine these indices together with the remaining free left-moving fermions, summed over Ramond and Neveu-Schwarz boundary conditions to obtain for the full trace

$$\begin{aligned}&\mathrm{Tr}_{\mathrm{R}} q^{L_0-c/24}\bar{q}^{\tilde{L}_0-\tilde{c}/24}J_0 e^{\pi i J_0}\\ &\quad = 2i\frac{Z_{10,2}}{\eta^{12}}\left[(\frac{\vartheta_3}{\eta})^{8-n}\Phi^{+}(\hat{A})-(\frac{\vartheta_4}{\eta})^{8-n}\Phi^{-}(\hat{A})+(\frac{\vartheta_2}{\eta})^{8-n}\Phi(\Delta)\right]\end{aligned}\tag{3.19}$$

where $Z_{10,2}/\eta^{12}(\tau)$ is the partition function for the free bosonic degrees of freedom on the $\Pi^{10,2}$ lattice constructed from the E_8 lattice and the $\Pi^{2,2}$ lattice of T^2.

Since the elliptic genus depends only on the topology of the manifold and the topology of the gauge bundle it is invariant under deformations of the hypermultiplet moduli in accord with the decompositions (2.2) and (2.1). Thus it can be evaluated by working in a special limit of the hypermultiplet moduli such as an orbifold limit. For the standard embedding of the spin connection in an $SU(2)$ subgroup of E_8 we can simply combine (3.19) with the results of [66] to obtain the answer.

4. Threshold corrections and the Prepotential

4.1. General strategy

In this section and the following we establish a connection between threshold corrections in $N=2$ compactifications of heterotic string theory and product formulae studied recently by Borcherds in connection with generalized Kac-Moody algebras. This strongly suggests the presence of a GKM algebra in such compactifications. In the ninth section we will discuss a construction of this GKM algebra in terms of vertex operators for BPS states.

We will determine the prepotential by comparing two formulae for one-loop renormalizations of nonabelian gauge couplings[6].

---

[6] In this section we follow the "string-theoretic" conventions of [48][38] for moduli. Thus, for example $\Re S>0$, $\Re T>0$, etc. The conventions which are useful for discussing automorphic properties are related by

$$y^{\mathrm{string}}=-iy^{\mathrm{automorphic}}$$

For a gauge group G the one-loop coupling renormalization is given by [48]:

$$\begin{aligned}\frac{1}{g^2(G;p^2)} &= k\Re\Big[S+\frac{1}{16\pi^2}\Delta^{\mathrm{univ}}\Big]+\frac{b(G)}{16\pi^2}\log\frac{M^2_{\mathrm{string}}}{p^2}+\frac{1}{16\pi^2}\Delta(G;y)\\ \Delta(G;y) &= \int_{\mathcal{F}}\frac{d^2\tau}{\tau_2}\Big[\mathcal{B}-b(G)\Big]\end{aligned}\tag{4.1}$$

where $\mathcal{B}$ is given by a trace over the internal Hilbert space:

$$\mathcal{B}=-\frac{i}{\eta^2}\mathrm{Tr}_{\mathcal{H}^{\mathrm{int}}_{\mathrm{R}}}\Big\{\mathrm{J}_0 \mathrm{e}^{\mathrm{i}\pi \mathrm{J}_0}\mathrm{q}^{\mathrm{L}_0-22/24}\bar{\mathrm{q}}^{\tilde{\mathrm{L}}_0-9/24}\big[\mathrm{Q}^2-\frac{\mathrm{k}}{8\pi\tau_2}\big]\Big\}\tag{4.2}$$

Here Q is a generator of the gauge group, $b(G)$ is the coefficient of the one-loop beta function normalized as in [48] and S is the dilaton field. The "universal" term Δ^{univ} is related to the "Green-Schwarz" term which governs the one-loop mixing of the axion and moduli fields and plays a role in the cancellation of sigma-model anomalies [67,68,48]. The quantity k is the level of the Kac-Moody algebra associated to G. Henceforth we set $k = 1$.

From the discussion in the previous section it is clear that (4.2) receives contributions only from BPS states. Morally speaking

$$\frac{1}{g^2}\sim\frac{1}{16\pi^2}\big[\sum_{vm}2Q^2\log m^2-\sum_{hm}2Q^2\log m^2\big]\tag{4.3}$$

in accord with the threshold rule of [38] eqn. 3.45.

We can further elaborate (4.1) using the constraints of $N = 2$ supergravity following the work of [38]. In (4.1) we may express the universal threshold correction in terms of the one-loop correction to the prepotential $h^{(1)}$:

$$\frac{1}{16\pi^2}\Delta^{\mathrm{univ}}=\frac{1}{-(\Re y)^2}\Re\Big[h^{(1)}-y_1^a\frac{\partial}{\partial y^a}h^{(1)}\Big]\tag{4.4}$$

On the other hand, using the Wilsonian coupling we may also write [48][38]:

$$\frac{1}{g^2(G;p^2)}=\Re\Big[\tilde{S}-\frac{1}{(s+4)\pi^2}\log[\Psi(y)]\Big]+\frac{b(G)}{16\pi^2}\Big[\log\frac{M^2_{\mathrm{Planck}}}{p^2}+K(S,\bar{S},y,\bar{y})\Big]\tag{4.5}$$

where

$$K=-\log[\Re(S)]-\log[-(y_1)^2]+const.\tag{4.6}$$

and a subscript 1 denotes the real part. The Planck mass is given by $M^2_{\mathrm{Planck}}=M^2_{\mathrm{string}}\Re S$. Since we are using the string theory conventions of [48,38] for the moduli the real part of y

is in the forward lightcone. The coupling $\tilde{S}$ differs from S by the addition of a holomorphic function such that $\tilde{S}$ is invariant under the duality group up to shifts. In [38] it is shown that we may write [7]:

$$\tilde{S} = S + \frac{1}{s+4}\eta^{ab}\frac{\partial}{\partial y^a}\frac{\partial}{\partial y^b}h^{(1)} \tag{4.7}$$

and under the perturbative duality group we have

$$\tilde{S} \to \tilde{S} + \frac{i}{2(s+4)}\eta^{IJ}\Lambda_{IJ} \tag{4.8}$$

where Λ_{IJ} is a real symmetric matrix.

In (4.5) $\Psi(y)$ is a holomorphic function on $\mathcal{H}^{s+1,1}$ such that

$$\Re\Big[\frac{1}{(s+4)\pi^2}\log\big[\Psi(y)\big] + \frac{b(G)}{16\pi^2}\log\big[-(y_1)^2\big]\Big] \tag{4.9}$$

is invariant under the perturbative duality group $O(s+2,2;\mathbb{Z})$. Requiring that the physical coupling be free of any singularity on $\mathcal{N}^{s+2,2}$ fixes the divisor of Ψ. These conditions determine $\Psi(y)$ up to an overall constant, since if we have two such then Ψ_1/Ψ_2 is a well-defined automorphic function extending to the compactification divisors of the variety $\mathcal{N}^{s+2,2}$ and is thus a constant. [8]

Equating (4.1) and (4.5) gives a formula for the prepotential. We now present the solution for some special cases. We will consider a compactification of the heterotic string on $K3 \times T^2$ with the standard embedding of the spin connection in the gauge group. This breaks the $E_8 \times E_8$ gauge group to $E_7 \times E_8$. In the further reduction to four dimensions on T^2 one can choose Wilson lines for the remaining unbroken subgroup leading at generic points to a low-energy gauge group $U(1)^{s+4}$ with $s = 15$ with the $U(1)^4$ arising from the two left-moving and two right-moving $U(1)$'s associated with the T^2 factor and the remaining $U(1)^{15}$ corresponding to the Cartan subalgebra of $E_7 \times E_8$.

When the unbroken gauge group is $[E_8 \times E_7 \times U(1)^2]_{left} \times [U(1)^2]_{right}$ (or an enhancement thereof) we will compute the running E_8 gauge coupling:

$$\frac{1}{g^2(E_8;p^2)} \qquad s = 0 \tag{4.10}$$

---

[7] Note the change of sign from 4.30 in [38]. We introduce $\tilde{S}$ so that we can also discuss the $(10,2)$ case. In the $(2,2)$ case it is related to the invariant coupling $S^{\rm inv}$ of [38] via: $\tilde{S} = S^{\rm inv} - L/8$.

[8] Strictly speaking one must specify appropriate asymptotic conditions to guarantee uniqueness. However we do not need the precise statement since we obtain the prepotential by direct computation.

as a function of the moduli T, U parametrizing $\mathcal{N}^{2,2}$. When the gauge group is $[E_7 \times U(1)^{10}]_{left} \times [U(1)^2]_{right}$ we will compute

$$\frac{1}{g^2(E_7; p^2)} \qquad s = 8 \tag{4.11}$$

as a function of the moduli in $\mathcal{N}^{10,2}$. The calculation for both cases can be done in parallel. The answer is given in section 4.4 below.

4.2. Computing the integrand

The first step in the calculation is the evaluation of the integrand in (4.1). As discussed in the previous section, the hypermultiplet dependence of the integrand enters only through the elliptic genus and therefore depends only on the topology of the gauge bundle and the manifold specifying the compactification. Thus, although the integrand in principle requires evaluating a partition function of a super conformal field theory on K3, we can perform the computation in an orbifold or other limit which is smoothly connected to the theory of interest. We will evaluate corrections by working in the T^4/Z_2 orbifold limit of $K3$ with the standard orbifold embedding of the twist in the gauge group. Related computations have been done for the elliptic genus of $K3$ in [66].

We start with the Narain lattice for a T^6 compactification to four dimensions with the decomposition

$$\Gamma^{22,6} = \Gamma^{10,2} \oplus \Gamma^{4,4} \oplus \Gamma^{8,0} \tag{4.12}$$

We then mod out by an involution which acts as -1 on the $\Gamma^{4,4}$ factor and as a shift $X^I \to X^I + \delta^I$ on $\Gamma^{8,0}$ with

$$\delta = (\tfrac{1}{2}, \tfrac{1}{2}, 0^6). \tag{4.13}$$

For this particular orbifold limit the unbroken gauge group at generic points is in fact $U(1)^{20}$ and the vectormultiplet moduli space is $\mathcal{N}^{18,2}$. However we will restrict ourselves in what follows to studying the dependence of the threshold corrections on the moduli of the $\Gamma^{10,2}$ factor in (4.12) for which the classical moduli space is $\mathcal{N}^{10,2}$. At least for orbifolds it should be possible to study the dependence on other moduli by generalizing the formulae in the appendix to congruence subgroups of $SL(2, Z)$.

We thus can evaluate:

$$\begin{aligned}
\eta^{-2}(\tau)\mathrm{Tr_R q^{L_0-c/24}\bar{q}^{\tilde{L}_0-\tilde{c}/24}J_0e^{\pi i J_0}} &= -2iZ_{10,2}\frac{E_6}{\eta^{24}} \\
\eta^{-2}(\tau)\mathrm{Tr_R q^{L_0-c/24}\bar{q}^{\tilde{L}_0-\tilde{c}/24}J_0e^{\pi i J_0}}[\mathrm{Q^2(E_7)}-\frac{1}{8\pi\tau_2}] &= -\frac{i}{12}Z_{10,2}\mathcal{E}_7 \\
\eta^{-2}(\tau)\mathrm{Tr_R q^{L_0-c/24}\bar{q}^{\tilde{L}_0-\tilde{c}/24}J_0e^{\pi i J_0}}[\mathrm{Q^2(E_8)}-\frac{1}{8\pi\tau_2}] &= -\frac{i}{12}Z_{2,2}\mathcal{E}_8
\end{aligned} \tag{4.14}$$

where

$$\begin{aligned}
Z_{s+2,2} &= \sum_{p\in\Gamma^{s+2,2}} q^{\frac{1}{2}p_L^2}\bar{q}^{\frac{1}{2}p_R^2} \\
\mathcal{E}_7 &= \left(\frac{(E_2-\frac{3}{\pi\tau_2})E_6-E_4^2}{\eta^{24}}\right) \\
\mathcal{E}_8 &= \left(\frac{(E_2-\frac{3}{\pi\tau_2})E_4E_6-E_6^2}{\eta^{24}}\right)
\end{aligned} \tag{4.15}$$

Here $E_{2n}(\tau)$ are the usual Eisenstein series and are given explicitly in the appendix. As a check note that the first equation in (4.14) follows from (3.19) and the results of [66].

It follows that we can rewrite the coupling in (4.1) as

$$\begin{aligned}
\frac{1}{g^2(G;p^2)} &= \Re\left[S+\frac{1}{16\pi^2}\Delta^{\mathrm{univ}}\right]+\frac{b(G)}{16\pi^2}\log\frac{M^2_{\mathrm{string}}}{p^2} \\
&\quad -\frac{1}{12}\frac{1}{16\pi^2}\left(\tilde{\mathcal{I}}_{s+2,2}-\mathcal{I}_{s+2,2}\right)
\end{aligned} \tag{4.16}$$

Here we have introduced a class of integrals defined in equations $(A.1),(A.2)$ of appendix A. The specific integrals appearing in (4.16) involve the modular forms:

$$\begin{aligned}
s=8:\qquad \frac{E_6}{\eta^{24}} &= \sum c_1(n)q^n = q^{-1}-480+\cdots \\
\frac{E_4^2}{\eta^{24}} &= \sum c_3(n)q^n = q^{-1}+504+\cdots
\end{aligned} \tag{4.17}$$

$$\begin{aligned}
s=0:\qquad \frac{E_6E_4}{\eta^{24}} &= \sum c_1(n)q^n = q^{-1}-240+\cdots \\
\frac{E_6^2}{\eta^{24}} &= \sum c_3(n)q^n = q^{-1}-984+\cdots
\end{aligned} \tag{4.18}$$

Note that the subtraction of the constant terms appearing in the definition of the integrals $\tilde{\mathcal{I}},\mathcal{I}$ and the beta function are consistent for

$$\begin{aligned}
b(G) &= -\frac{1}{12}(\tilde{c}_1(0)-c_3(0)) \\
&= -60 \qquad s=0 \\
&= +84 \qquad s=8
\end{aligned} \tag{4.19}$$

In general the coefficients $c_1(n)$ and $c_3(n)$ depend both on the choice of gauge bundle and on which low-energy gauge group is being studied.

4.3. Formulae for the Integrals

The integrals $\mathcal{I}, \tilde{\mathcal{I}}$ are evaluated in appendix A. The integral $\mathcal{I}$ is given by the expression:

$$\begin{aligned}\mathcal{I}_{s+2,2}(y) &= -2\log\Big|\Phi(y)\Big|^2 + c_3(0)\Big(-\log[-(\Re y)^2] - \mathcal{K}\Big)\\ \Phi(y) &= e^{-2\pi\rho\cdot y}\prod_{r>0}\Big(1-e^{-2\pi r\cdot y}\Big)^{c_3(-r^2/2)}\end{aligned} \tag{4.20}$$

In the above $\mathcal{K}$ is a constant defined in the appendix. The product over $r > 0$ means the following. We consider the even-self-dual Lorentzian lattice $\Pi^{s+1,1}$ for $s = 0, 8$ and write lattice vectors as $r = (\vec{b}, -\ell, -k)$ with $\vec{b} \in \Pi^{s,0}$. For $s = 8$, $\Pi^{8,0}$ is the root lattice of E_8 and we choose a set of positive roots for this lattice. With this understood, $r > 0$ means:

1. $k > 0$ or,
2. $k = 0$, $\ell > 0$, or,
3. $k = \ell = 0$, $\vec{b} > 0$.

For $s = 8$ we show in section seven below that this is the positive root condition for E_{10}. For $s = 0$ the coefficient $c_3(0)$ is not determined purely by modular invariance. After explicitly dividing out the terms involving $c_3(0)$ the remaining product is over the positive roots of the Monster Lie algebra [34][35].

The vector ρ is:

$$\begin{aligned}\rho = &\rho_{E_{10}} = -(\vec{\rho}; 31, 30) \qquad s = 8\\ \rho = &-\frac{c_3(0)}{24}(1,1) + (0,1) \qquad s = 0.\end{aligned} \tag{4.21}$$

where $\vec{\rho}$ is the Weyl vector of E_8. For $s = 8$, $\rho_{E_{10}}$ is the Weyl vector for E_{10}. For $s = 0$ (4.21) gives a lattice Weyl vector for the $\Pi^{1,1}$ lattice but only gives the Weyl vector for the Monster Lie algebra after subtraction of the $c_3(0)$ term.

The second integral we need is[9]:

$$\begin{aligned}\tilde{\mathcal{I}}_{s+2,2}(y) = &4\Re\Bigg\{\sum_{r>0}\Big[\tilde{c}_1(-\frac{r^2}{2})Li_1(e^{-2\pi r\cdot y}) + \frac{6}{\pi(y_1)^2}c_1(-\frac{r^2}{2})\mathcal{P}(ir\cdot y)\Big]\Bigg\}\\ &+\tilde{c}_1(0)\Big(-\log[-(y_1)^2] - \mathcal{K}\Big) + \frac{1}{(y_1)^2}[\tilde{d}^{s+2,2}_{ABC}y_1^A y_1^B y_1^C + \delta]\end{aligned} \tag{4.22}$$

---

[9] Integrals of this type are also evaluated in [69] although there are important differences in our results.

where

$$\delta = \frac{6}{\pi^2} c_1(0)\zeta(3). \tag{4.23}$$

Here $Li_1(x) = -\log(1-x)$ and the function $\mathcal{P}$ involves the polylogarithms Li_2 and Li_3 and is defined in the appendix. As explained in section 5.1, $\mathcal{H}^{s+1,1}$ is tessellated by a system of Weyl chambers, and the symmetric tensor $\tilde{d}^{s+2,2}$ is piecewise constant in $\mathcal{H}^{s+1,1}$, taking different values in different Weyl chambers. Explicit formulae for it are given in appendix A.

4.4. Answer for the prepotential

Equating (4.16) with (4.5) and using (4.4) and (4.7) we obtain a differential equation for $h^{(1)}$:

$$\begin{aligned} &\Re\Big[\frac{1}{s+4}\eta^{ab}\frac{\partial}{\partial y^a}\frac{\partial}{\partial y^b}h^{(1)}\Big] + \frac{1}{(\Re y)^2}\Re\Big[h^{(1)} - y_1^a\frac{\partial}{\partial y^a}h^{(1)}\Big] \\ &\qquad = -\frac{1}{12}\frac{1}{16\pi^2}\Big(\tilde{\mathcal{I}}_{s+2,2} - \mathcal{I}_{s+2,2}\Big) + \frac{1}{(s+4)\pi^2}\Re\big[\log\Psi\big] + \frac{b(G)}{16\pi^2}\log(-y_1^2). \end{aligned} \tag{4.24}$$

A particular solution of the differential equation (4.24), *in the fundamental Weyl chamber*, [10] is given by

$$h^{(1)} = \frac{1}{384\pi^2}\tilde{d}^{s+2,2}_{ABC}y^A y^B y^C - \frac{1}{2(2\pi)^4}c_1(0)\zeta(3) - \frac{1}{(2\pi)^4}\sum_{r>0} c_1(-r^2/2)Li_3(e^{-2\pi r\cdot y}) \tag{4.25}$$

Substitution into (4.24) leads to a solution with

$$\log\Psi(y) = \tfrac{1}{2}\log[J(iT) - J(iU)] + b(E_8)\log[\eta(iT)\eta(iU)] \qquad s = 0 \tag{4.26}$$

and

$$\Psi(y) = \Phi(y) = e^{-2\pi\rho\cdot y}\prod_{r>0}\Big(1 - e^{-2\pi r\cdot y}\Big)^{c_3(-r^2/2)} \qquad s = 8. \tag{4.27}$$

Proof. We substitute directly the ansatz for $h^{(1)}$. The universal term is given by

$$\frac{1}{16\pi^2}\Delta^{\text{univ}} = \frac{1}{y_1^2}\frac{1}{(2\pi)^3}\Re\Big[\sum_{r>0} c_1(-r^2/2)\mathcal{P}(ir\cdot y)\Big] + \frac{1}{192\pi^2(y_1)^2}[\tilde{d}^{s+2,2}_{ABC}y_1^A y_1^B y_1^C + \delta] \tag{4.28}$$

---

[10] See section 5.1

In order to cancel the other terms we must use the identity:

$$
\begin{aligned}
\sum n c_1(n) q^n &= -\frac{1}{2}\frac{E_2E_6+E_4^2}{\eta^{24}} \qquad\qquad s=8 \\
\sum n c_1(n) q^n &= -\frac{1}{6}\frac{E_2E_4E_6+2E_6^2+3E_4^3}{\eta^{24}} \qquad s=0
\end{aligned}
\tag{4.29}
$$

One can check the rational terms directly, but a more elegant method relates the trace of $\tilde{d}_{ABC}$ to the Weyl vector. For example, in the case $s=8$ we proceed as follows. The Laplacian for the Kahler metric on $\mathcal{N}^{s+2,2}$ is given by

$$
\nabla^2 = -2y_1^2\Big(\eta^{ab} - \frac{2}{y_1^2}y_1^a y_1^b\Big)\partial_a\bar{\partial}_{\bar{b}} \tag{4.30}
$$

By straightforward computation one finds:

$$
\begin{aligned}
(\nabla^2-24)\tilde{\mathcal{I}}_{10,2} + 280\tilde{c}_1(0) = 24\Big[&-2\log|\Phi(y)|^2 \\
&- c_3(0)(\log(-y_1^2)+\mathcal{K}) + (8\pi\rho_b - \frac{1}{4}\tilde{d}^a{}_{ab})y_1^b\Big]
\end{aligned}
\tag{4.31}
$$

since both sides of the equation must be invariant under the duality group we have after comparing to (4.20):

$$
\tilde{d}^a{}_{ab} = -32\pi\rho_b \tag{4.32}
$$

Similarly one finds

$$
\begin{aligned}
(\nabla^2-4)\tilde{\mathcal{I}}_{2,2} &= 4\tilde{c}(0)\big(\log\big[2T_1U_1|\eta(iT)^4\eta(iU)^4|\big] + \mathcal{K} - \tfrac{1}{2}\big) \\
&\quad - 80\log|J(iT)-J(iU)|
\end{aligned}
\tag{4.33}
$$

Using this equation one can check that the rational terms in (4.25) solve the equation required for equality of (4.1) and (4.5).

Above we have exhibited a particular solution of the second order differential equation. Two solutions to this equation must differ by a solution of the homogeneous equation. It is straightforward to show that the only solutions of the homogeneous equation which are analytic around zero are of the form:

$$
\delta h^{(1)} = \sum_{0\le m+n\le 2} a_{IJ,mn}(\hat{X}^I)^m(\hat{X}^J)^n \tag{4.34}
$$

where the coefficients $a_{IJ,mn}$ are pure imaginary and [11]

$$\hat{X}^I = (1, -y^2, iy^a). \tag{4.35}$$

Two prepotentials differing by such an expression are physically equivalent. We may use this result, together with the automorphic properties discussed in section six to argue that the prepotential given above is the unique answer.

It is worthwhile to make several checks on the above answer. First, the physical coupling must be a nonsingular function on moduli space. We can check this since the singularities of $\log \Psi$ are identical to $\frac{1}{s+4}\left(\frac{\partial}{\partial y}\right)^2 h^{(1)}$. Second, the physical coupling must be duality invariant. This is a consequence of the automorphic properties proved in the next two sections. Third, in the case $s = 0$, a differential equation for the quantity $S^{\rm inv} = S - \frac{1}{16\pi^2}\Delta^{\rm univ}$ was derived in [48,38]. [12] It is straightforward to show that the above formula for $h^{(1)}$ satisfies this constraint. As a fourth check we may compare the "Yukawa couplings" [13] following from (4.25) with those derived in [38][39]. For example, one finds, in the $s = 0$ case:

$$\partial_U^3 h^{(1)} = -\frac{1}{2\pi}\left[1 - \sum_{r>0} c_1(k\ell)\ell^3 \frac{e^{-2\pi(kT+\ell U)}}{1 - e^{-2\pi(kT+\ell U)}}\right] \tag{4.36}$$

According to [38,39] (4.36) must coincide with

$$-\frac{1}{2\pi}\frac{E_4(iU)E_4(iT)E_6(iT)}{(J(iT) - J(iU))\eta^{24}(iT)} \quad . \tag{4.37}$$

Comparing the $T \to \infty$ and $U \to \infty$ limits of the expressions one finds perfect agreement. Expanding in power series and comparing terms we find agreement to 10^{th} order.

4.5. Gravitational Corrections

There are other terms in the effective supergravity action which involve chiral densities and are therefore computable to all orders of perturbation theory. The first of these is the gravitational coupling F_1 given by

$$F_1 = \frac{-i}{192\pi^2}\int_{\mathcal{F}} \frac{d^2\tau}{\tau_2}\left[\frac{1}{\eta^2}{\rm Tr}_{\rm R} {\rm J}_0 (-1)^{{\rm J}_0} {\rm q}^{{\rm L}_0 - 22/24}\bar{\rm q}^{\tilde{\rm L}_0 - 9/24}\left[{\rm E}_2 - \frac{3}{\pi\tau_2}\right] - {\rm b}_{\rm grav}\right] \tag{4.38}$$

---

[11] in a gauge where $X^0 = 1$

[12] See, for example, equation 4.46 in [38].

[13] Actually, magnetic moments, [60].

where the E_2 factor arises from the $Q^2_{grav} = -2\partial_\tau \log \eta$ term in [43]. In heterotic string theory F_1 and its generalizations F_g can be computed at one-loop order and the comparison between these calculations and genus g computations in dual twisted Calabi-Yau theories provide strong evidence for $N=2$ string duality [58,56]. Again it is clear from the previous discussion that these quantities in heterotic string theory receive contributions only from BPS states. There are also potential ambiguities in these amplitudes which require a careful treatment of infrared divergences, perhaps along the lines presented in [69] [14].

5. Comment on relation to work of Borcherds

In this section we will discuss the automorphic properties of the threshold corrections we have computed and compare our results to work of Borcherds. We will work in "automorphic" conventions for the moduli as described in footnote 4 of the previous section.

First we explain the relation between Borcherds' "rational quadratic divisors" (RQD) and enhanced symmetry points (ESP) in Narain compactifications. Let $L \cong \Pi^{s+1,1}$ be an even unimodular lattice, and consider the $\Pi^{s+2,2}$ lattice $M \cong L \oplus \Pi^{1,1}$. We represent an element of M by $v = (r; a_+, a_-)$ with $r \in L$ and $a_\pm$ integer. In [29] Borcherds defines rational quadratic divisors to be the locus $\mathcal{D}(v) \subset \mathcal{H}^{s+1,1}$ of

$$\langle (r; -a_+, a_-), (y; 1, -\tfrac{1}{2}(y,y)) \rangle = 0 \tag{5.1}$$

for

$$\langle (r; -a_+, a_-), (r; -a_+, a_-) \rangle = r^2 - 2a_+a_- > 0 \tag{5.2}$$

From the discussion in sec. 3.1 we know that BPS states are parametrized by vectors $v = (r; a_+, a_-) \in \Pi^{s+2,2} \cong \Pi^{s+1,1} \oplus \Pi^{1,1}$ The central charge of a BPS state with quantum numbers $(r; a_+, a_-) \in \Pi^{s+1,1} \oplus \Pi^{1,1}$ is just the inner product [37]

$$\mathcal{Z}(v, y) = (r, y) + a_- - \tfrac{1}{2} a_+ y^2 = \langle (r; a_+, a_-), (y; 1, -\tfrac{1}{2} y^2) \rangle \tag{5.3}$$

Moreover, for such a state

$$p_L^2 - p_R^2 = r^2 + 2a_+a_- \tag{5.4}$$

---

[14] We thank J. Louis for pointing this out to us.

Thus in string theory only the divisors with $\langle (r;-a_+,a_-),(r;-a_+,a_-)\rangle = 2$ are of importance and these correspond to enhanced symmetry points in the Narain moduli space.

In [29][30] Borcherds has proved the following theorem:

Theorem. Let $f(\tau)=\sum c(n)q^n$ be a meromorphic modular form with all poles at cusps. Suppose that f is of weight $-s/2$ for $SL(2,\mathbb{Z})$ and has integer coefficients, with $24|c(0)$ if $s=0$. Then there is a unique vector $\rho\in L$ such that

$$\Phi(y)=e^{2\pi i\rho\cdot y}\prod_{r>0,r\in\Pi^{s+1,1}}(1-e^{2\pi i r\cdot y})^{c(-r.r/2)} \tag{5.5}$$

can be analytically continued to define a meromorphic automorphic form of weight $c(0)/2$ for $O(s+2,2;\mathbb{Z})^+$. All the zeroes and poles of Φ lie on rational quadratic divisors, and the multiplicity of the zero of Φ at the rational quadratic divisor of the triple $(b;a_1,a_2)$ is

$$\sum_{n>0}c(n^2(a_1a_2-(b,b)/2)) \tag{5.6}$$

In (5.5) $r>0$ means that r has positive inner product with a chosen negative norm vector in L. In some cases the product (5.5) is known to be the denominator formula for a generalized Kac-Moody algebra.

We have seen that products of precisely this type arise in the analysis of threshold corrections in $N=2$ heterotic string compactifications. Indeed, by studying the properties of the integrals $\mathcal{I}_{s+2,2}(y)$ of appendix A we can easily rederive many of the results discovered by Borcherds.

To do this we must explain the holomorphic factorization of the product arising in the expression for $\mathcal{I}$ in equation $(A.29)$ of appendix A. In the previous section we presented the results only in a particular Weyl chamber (defined below). However the integrals as evaluated in the appendix depend on the "hatted" dot product $r\hat{\cdot}y$ rather than the dot product $r\cdot y$. The quantity $r\hat{\cdot}y$ is defined by:

$$r\,\hat{\cdot}\,y\equiv\Re\Big[(\vec{b}\cdot\vec{y}+\ell y_-+ky_+)\Big]+i\Big|\Im\Big[(\vec{b}\cdot\vec{y}+\ell y_-+ky_+)\Big]\Big| \tag{5.7}$$

when $k>0$, and, when $k=0$:

$$r\,\hat{\cdot}\,y\equiv r\cdot y-N(r,y)y_- \tag{5.8}$$

where

$$N(r,y)=\mathrm{sgn}(\vec{b})\left[\mathrm{sgn}(\vec{b})\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}\right] \tag{5.9}$$

is an integer ($[\cdot]$ is the greatest integer function.) In order to understand the holomorphic factorization we must introduce Weyl chambers.

5.1. Weyl chambers

We can tessellate the forward lightcone $C_{+}^{s+1,1} \subset \mathcal{H}^{s+1,1}$ into convex polyhedra whose walls are defined by the *real* codimension one subvarieties

$$\Im(r \cdot y) = 0 \tag{5.10}$$

for $r^2 = 2, r \in \Pi^{s+1,1}$. These are the "surfaces of marginal stability" and must be crossed when circling the divisors of vanishing BPS central charges.

Given a choice of simple roots r_μ for the lattice $\Pi^{s+1,1}$ we can define a *fundamental Weyl chamber* by the equation

$$\mathcal{CW}_0 \equiv \{y : r_\mu \cdot \Im y > 0\} \tag{5.11}$$

Explicit choices of simple roots are described in section seven. For $s = 0$ there are two Weyl chambers and we take the fundamental Weyl chamber to be simply $y_{+,2} > y_{-,2}$. For $s = 8$ we choose a set of positive roots $\vec{b} > 0$, $\vec{b}^2 = 2$ for E_8. Using the simple roots in the next section we get the conditions:

$$\begin{aligned} 0 < & \frac{\vec{b} \cdot \Im \vec{y}}{y_{-,2}} < 1 \\ y_{-,2} & < y_{+,2} \end{aligned} \tag{5.12}$$

5.2. Holomorphic factorization and $\mathcal{I}_{s+2,2}(y)$

By definition of the fundamental Weyl chamber we know that $r \hat{\cdot} y = r \cdot y$ for the positive roots. Thus, the products $(A.31), (A.37)$ factorize straightforwardly in the fundamental chamber. The question arises as to the other chambers.

Contributions from $c(n)$ for $n > 0$ always holomorphically factorize. To see this note that $\Im y$ is in the forward light cone so

$$\begin{aligned} \Im y_+ &> 0 \\ \Im y_- &> 0 \\ 2 \Im y_+ \Im y_- &> (\Im \vec{y})^2 \end{aligned} \tag{5.13}$$

Since we are studying coefficients $c(n)$ for $n > 0$ in the product we have $2k\ell > (\vec{b})^2$. It follows that

$$(k \Im y_+ + \ell \Im y_-)^2 > 4(k \Im y_+)(\ell \Im y_-) > |\vec{b} \cdot \Im \vec{y}|^2 \tag{5.14}$$

so the imaginary part of $r \cdot y$ is always positive and the product is a holomorphic square. Similarly, the coefficient $c(0)$ enters for roots $k = 0, \ell > 0, \vec{b} = 0$ and for roots $k\ell = \frac{1}{2}\vec{b}^2 > 0$. For these roots we can again drop the hat. All of this is quite simply understood from the physical point of view: only states with $v^2 = 2$ can become massless and lead to nonanalytic behavior.

The real source of nontrivial holomorphic factorization is entirely in the coefficients of $c(-1)$ which are connected to RQD's or ESP's. For example, the peculiar shift in (5.9) affects these terms. The key observation is that these changes can be absorbed in a change of the linear term $\rho \cdot y \to \rho_\alpha \cdot y$ in the RHS of $(A.29)$ in each Weyl chamber $\mathcal{C}_\alpha$. Consider for example, the holomorphic factorization for roots with $r^2 = +2$ and $k > 0$. Then:

$$\log\left|1 - e^{2\pi i \widehat{r \cdot y}}\right|^2 = \begin{cases} \log\left|1 - e^{2\pi i r \cdot y}\right|^2 & \text{for } \Im r \cdot y > 0 \\ \log\left|e^{-2\pi i r \cdot y}(1 - e^{2\pi i r \cdot y})\right|^2 & \text{for } \Im r \cdot y < 0 \end{cases} \tag{5.15}$$

For example, consider the case $s = 0$. There is only one positive root with $r^2 = 2$, namely $r = (\ell, k) = (-1, 1)$. The wall is just

$$\Im y_+ = \Im y_- \tag{5.16}$$

i.e. $T_2 = U_2$. Consequently the product in (4.20) becomes

$$-2 \log\left| e^{-2\pi i T c(-1)} \prod_{r>0} \left(1 - e^{2\pi i (kT + \ell U)}\right)^{c(k\ell)} \right|^2 \tag{5.17}$$

for $T_2 > U_2$ and

$$-2 \log\left| e^{-2\pi i U c(-1)} \prod_{r>0} \left(1 - e^{2\pi i (kU + \ell T)}\right)^{c(k\ell)} \right|^2 \tag{5.18}$$

for $U_2 > T_2$.

5.3. Automorphic properties of Φ

We will now use invariance of the integral under the group generated by

1. $y \to y + \lambda \ , \ \lambda \in \Pi^{s+1,1}$
2. $y \to w(y)$, $w \in O(s+1, 1; \mathbb{Z})^+$
3. $y \to y' = 2y/(y, y)$

to deduce some automorphic properties of the "holomorphic square root" $\Phi(y)$. It is evident that the first two transformations act on $\mathcal{H}^{s+1,1}$. The third is surprising, but note that under this transformation:

$$(\Im y', \Im y') = +\frac{4}{|(y,y)|^2}(\Im y, \Im y) \tag{5.19}$$

and so preserves the "upper half plane." In fact, the transformations 1.,2. and 3. only generate an index 2 subgroup of the duality group, and Φ is automorphic for the entire group $O(s+2,2;\mathbb{Z})$.

Invariance under $y \to y + \lambda$ is trivial. To verify $y \to w(y)$ invariance we use (5.15) which says that in a Weyl chamber $\mathcal{C}_\alpha$ we have

$$\mathcal{I} = -2\log|e^{(\rho_\alpha - \rho)\cdot y}\Phi(y)|^2 + c(0)\Big(-\log[-(\Im y)^2] - \mathcal{K}\Big) \tag{5.20}$$

Now we note that

$$(\rho_\alpha - \rho) = w(\rho) - \rho \tag{5.21}$$

To prove this note that it is true for chambers related to the fundamental chamber by a reflection in a simple root r_i. Then proceed from there by induction.

The transformation of Φ under the third generator is the most difficult to prove by straightforward methods. We may note that if $\Re y = 0$ then the transformation is a rescaling of $\Im y$ by a positive coefficient. Therefore, an open set of the fundamental Weyl chamber is mapped to itself. From the invariance of the integral $\mathcal{I}_{s+2,2}$ and its holomorphic factorization we then find the modular transformation law:

$$\Phi(y') = \Big[\frac{(y,y)}{2}\Big]^{c(0)/2}\Phi(y) \tag{5.22}$$

5.4. Zeroes and Poles of Φ

Finally, we come to the last part of Borcherds' theorem which is concerned with the zeros and poles of Φ. Unlike the walls of the Weyl chambers, which are determined by $r \in \Pi^{s+1,1}$, the RQD's are determined by vectors $v \in \Pi^{s+2,2}$ with $v^2 = 2$. As $y \to \mathcal{D}(v)$ for some vectors in the Narain lattice $\Gamma^{s+2,2}(y)$, $p_R \to 0$, $p_L^2 \to 2$. Each such vector contributes a logarithmic divergence to $\mathcal{I}_{r+2,2}(y)$:

$$\int^{\infty} \frac{d\tau_2}{\tau_2} e^{-4\pi\tau_2 \frac{1}{2}p_R^2} \to -\log|\mathcal{Z}(v,y)|^2 \tag{5.23}$$

Analysis of RQD's: (2,2) case

The RQD's in $\mathcal{H}^{1,1}$ are given by the divisor $T = U$ and modular transformations of it. These have a double intersection at the points $T = U = i$ and a triple intersection at $T = U = \rho$ (and their modular images). This is in accord with standard analysis of the enhanced symmetry points.

One proof of the above statements proceeds by writing:

$$\mathcal{Z} = \begin{pmatrix} y_+ & 1 \end{pmatrix} \begin{pmatrix} n_0 & n_- \\ m_+ & m_0 \end{pmatrix} \begin{pmatrix} y_- \\ 1 \end{pmatrix} \tag{5.24}$$

where $r^2 = 2$ implies $m_0 n_0 - m_+ n_- = 1$. Note that $SL(2,Z) \times SL(2,Z)$ acts on the sets of divisors. By acting on y_- we can transform the matrix to the unit matrix, so the divisor equation becomes $1 + y_+ y_- = 0$ or, by a further $SL(2,\mathbb{Z})$ transformation $y_+ = y_-$. We obtain the set of all divisors in $\mathcal{H} \times \mathcal{H}$ by taking $SL(2,Z) \times SL(2,Z)$ images of this divisor. Note that divisors can only intersect when both y_+ and y_- are fixed by some element of $SL(2,\mathbb{Z})$. In particular, the divisor $y_+ = y_-$ has a double intersection at i and a triple intersection at ρ.

Analysis of RQD's: (10,2) case

It would be interesting to give a classification of the RQD's in $\mathcal{H}^{9,1}$ analogous to that above. An important class of RQD's are defined by the simple roots of E_8. For example, define the hyperplane

$$\mathcal{H}_i = \{y : \vec{y} \cdot \vec{\alpha}_i = 0\} \tag{5.25}$$

Along this hyperplane Φ has a simple zero

$$\Phi(y) \rightarrow -2\pi i \vec{\alpha}_i \cdot \vec{y} \Phi'(y') \tag{5.26}$$

where, for $y' \in \mathcal{H}_i$ we have

$$\Phi'(y') = e^{2\pi i \rho' \cdot y'} \prod_{r' > 0}{}' (1 - e^{2\pi i r' \cdot y'})^{c_3'(-r^2/2)} \tag{5.27}$$

where $r' \sim r$ if $r - r' \propto \alpha_i$ and $c_3'(-r^2/2) = \sum_{r' \sim r} c_3(-(r')^2/2)$. Restricting to successive intersections of divisors we get a series of automorphic forms beginning with the E_{10} form Φ of (4.20) and ending with $J(T) - J(U)$.

6. Quantum Monodromy

6.1. General Remarks

The nature of the semiclassical monodromy of $N=2$ heterotic string compactifications has been thoroughly analyzed in references [38,39]. Using the above explicit expressions we can compute the monodromy directly.

As discussed in [60,38,39] the best basis of special coordinates for questions of monodromy is the basis $\hat{X}^I$ related to the basis S, y^a with prepotential $\mathcal{F}=(X^0)^2 S(y,y)$ by a duality transformation $\hat{X}^1 = F_1, \hat{F}_1 = -X^1$. Explicitly we have:

$$\begin{aligned}\hat{X}^1 &= X^0 (y,y)^2 \\ \hat{X}^a &= X^0 y^a \qquad a = 2, \dots s+3\end{aligned} \tag{6.1}$$

In this basis the general transformation rule for the one-loop prepotential, in automorphic conventions is:

$$h^{(1)}(\tilde{y}) = \left(\frac{\tilde{X}^0}{X^0}\right)^{-2} \left[h^{(1)}(y) - i\Lambda_{IJ}\frac{\hat{X}^I}{X^0}\frac{\hat{X}^J}{X^0}\right] \tag{6.2}$$

where Λ_{IJ} is a real symmetric matrix. Thus, the prepotential is an automorphic form of weight -2 transforming with shifts.

In discussing monodromy one should always bear in mind that the prepotential is ambiguous by an addition of a term of the form (4.34).

6.2. Monodromies for $s=0$

It is convenient to introduce the function

$$\begin{aligned}\mathcal{L}(T,U) &\equiv \sum_{r>0} c(-r^2/2) Li_3(e^{2\pi i(kT+\ell U)}) \\ &= Li_3(e^{2\pi i(T-U)}) + \sum_{k,\ell\geq 0} c(k\ell) Li_3(e^{2\pi i(kT+\ell U)})\end{aligned} \tag{6.3}$$

where

$$F(q) = \sum_{n=-1}^{\infty} c(n) q^n = \frac{E_4 E_6}{\eta^{24}} \quad . \tag{6.4}$$

This function has a branch locus at $T=U$ and can be defined by power series for $T_2 > U_2$, i.e., in the fundamental Weyl chamber.

The function Li_3 can be analytically continued outside the unit circle and satisfies the connection formula [70]:

$$Li_3(e^x) = Li_3(e^{-x}) + \frac{\pi^2}{3}x - \frac{i\pi}{2}x^2 - \frac{1}{6}x^3 \tag{6.5}$$

Thus, under analytic continuation into the other Weyl chamber with $U_2 > T_2$ we have:

$$\mathcal{L}(T,U) = \mathcal{L}(U,T) - (2\pi)^4 \frac{i}{24\pi}\left[2U^3 - 2T^3 - 3(T-U)^2 + (T-U)(6TU-1)\right] \tag{6.6}$$

Under the generators $T \to T+1$ and $U \to U+1$ of the duality group the function $\mathcal{L}(T,U)$ is invariant in its region of convergence. Finally, to compute the monodromy under $T \to -1/T$ (and similarly for $U \to -1/U$) we rewrite the integral $\tilde{\mathcal{I}}_{2,2}$ as:

$$\begin{aligned} \tilde{\mathcal{I}}_{2,2} = -\frac{6}{U_2 T_2 \pi^2}\Re\left[\left(1 - iU_2\frac{\partial}{\partial U}\right)\left(1 - iT_2\frac{\partial}{\partial T}\right)\mathcal{L}(T,U)\right] \\ +20\log|J(T)-J(U)| + 264[\log[2T_2U_2|\eta(T)\eta(U)|^4] - \mathcal{K}] - \frac{\delta}{2T_2U_2} \\ +16\pi\frac{U_2^2}{T_2} \qquad T_2 \geq U_2 \\ +16\pi\frac{T_2^2}{U_2} \qquad U_2 \geq T_2 \end{aligned} \tag{6.7}$$

We know the integral $\tilde{\mathcal{I}}_{2,2}$ is invariant under $T \to T' = -1/T, U \to U$ and from this is it straightforward to derive the transformation law:

$$\begin{aligned} \mathcal{L}(T',U) &= T^{-2}\Big[\mathcal{L}(T,U) + \mathcal{R}(T,U)\Big] \\ \mathcal{R}(T,U) &= \frac{\pi^2}{12}\delta(1-T^2) - \frac{4\pi^3 i}{3}U^3(T^2-1) \qquad T_2' > U_2 \\ &= \frac{\pi^2}{12}\delta(1-T^2) - \frac{4\pi^3 i}{3}\left(\frac{1-T^4}{T}\right) + \frac{4\pi^3 i}{3}U^3 \qquad T_2' < U_2 \end{aligned} \tag{6.8}$$

In the 2,2 case there are two Weyl chambers, and accordingly we have the two expressions for the prepotential (in automorphic conventions):

$$\begin{aligned} h^{(1)} &= -\frac{1}{(2\pi)^4}\mathcal{L}(T,U) - \frac{\delta}{192\pi^2} - i\frac{U^3}{12\pi} + \delta h_+^{(1)} \qquad T_2 > U_2 \\ &= -\frac{1}{(2\pi)^4}\mathcal{L}(U,T) - \frac{\delta}{192\pi^2} - i\frac{T^3}{12\pi} + \delta h_-^{(1)} \qquad T_2 < U_2 \end{aligned} \tag{6.9}$$

where $\delta h_\pm^{(1)}$ correspond to the ambiguity (4.34).

Using these expressions and the above transformation rules for $\mathcal{L}(T,U)$ it is straightforward to compute the monodromy of $h^{(1)}$.

6.3. Remarks on the $s = 8$ case

In principle the above discussion generalizes in a fairly straightforward way to the computation of the monodromy on $\mathcal{N}^{10,2}$. The Weyl reflections σ_i in the E_{10} roots no longer square to zero because of the monodromy associated with the Li_3 terms around enhanced symmetry varieties. The transformation of $\mathcal{L}(y)$ under the inversion $y \to 2y/(y,y)$ can be deduced from an analog of (6.7) for $\tilde{\mathcal{I}}_{10,2}$. We hope to return to a detailed discussion of this monodromy representation, which gives an interesting representation of the E_{10} braid group, in the future.

7. GKM algebras and denominator formulae

In this section we will provide a very brief summary of some of the properties of GKM algebras and the hyperbolic Kac-Moody algebra E_{10} which enter into the interpretation of the product formulae for threshold corrections found in the previous section. For further details on Kac-Moody algebras and GKM algebras the reader can consult [34,71,72,73,25]. E_{10} is discussed in [74,75,26,76,28]

We first give the formal definition in terms of Cartan matrices, generators and relations. Recall the usual definition of a finite-dimensional simple Lie algebra. One starts with a symmetric $r \times r$ Cartan matrix $A = (a_{ij})$ with $i, j = 1, 2, \cdots r$ satisfying

$$\begin{aligned} &a_{ii} = 2, \\ &a_{ij} \le 0 \quad \text{for} \quad i \ne j, \\ &a_{ij} \in Z, \end{aligned} \tag{7.1}$$

and $\det A > 0$. The Lie algebra is then defined in terms of generators (e_i, f_i, h_i) obeying the relations

$$\begin{aligned} [h_i, h_j] &= 0, \\ [h_i, e_j] &= a_{ji} e_j, \\ [h_i, f_j] &= -a_{ji} f_j, \\ [e_i, f_j] &= \delta_{ij} h_j \\ (\mathrm{ad}_{e_i})^{1-a_{ji}} e_j &= 0, \\ (\mathrm{ad}_{f_i})^{1-a_{ji}} f_j &= 0. \end{aligned} \tag{7.2}$$

A formal construction of Kac-Moody Lie algebras is obtained by weakening the condition that $\det A > 0$. In particular one obtains affine Lie algebras by allowing A to have a zero eigenvalue and hyperbolic Kac-Moody algebras by allowing a single negative eigenvalue. For Kac-Moody algebras one has the usual notion of root spaces, positive roots, simple roots, and the Weyl group generated by reflections in the real simple roots. Furthermore there is a denominator formula

$$e^{\rho}\prod_{r>0}(1-e^{r})^{\mathrm{mult}(r)}=\sum_{\sigma\in W}(\mathrm{sgn}(\sigma))e^{\sigma(\rho)} \tag{7.3}$$

where ρ is the Weyl vector, the product on the left hand side runs over all positive roots and each term is weighted by the root multiplicity $\mathrm{mult}(r)$, and on the right hand side the sum is over all elements of the Weyl group W. For example for the $su(2)$ level one affine Lie algebra (7.3) is just the Jacobi triple product identity.

Generalized Kac-Moody algebras have a great deal in common with Kac-Moody algebras of hyperbolic type but differ from them in that simple roots r with $r^2 \leq 0$ are allowed. The formal definition follows from a slight generalization of the conditions (7.1). One again starts with a symmetric Cartan matrix $A=(a_{ij})$ which is allowed to have infinite rank in general. Then one demands the conditions

$$\begin{aligned} a_{ii}&=2 \quad \text{or} \quad a_{ii}\leq 0,\\ a_{ij}&\leq 0 \quad \text{for} \quad i\neq j,\\ a_{ij}&\in Z \quad \text{if} \quad a_{ii}=2, \end{aligned} \tag{7.4}$$

and defines the algebra by generators (h_{ij}, e_i, f_i) and relations

$$\begin{aligned} [h_{ij},e_k]&=\delta_{ij}a_{jk}e_k,\\ [h_{ij},f_k]&=-\delta_{ij}a_{jk}f_k,\\ [e_i,f_j]&=h_{ij}\\ (\mathrm{ad}_{e_i})^{1-a_{ji}}e^{j}&=0, \quad (\mathrm{ad}_{f_i})^{1-a_{ji}}f^{j}=0, \quad a_{ii}=2, i\neq j,\\ [e_i,e_j]&=0, \quad [f_i,f_j]=0, \quad a_{ii}\leq 0, a_{jj}<0, a_{ij}=0. \end{aligned} \tag{7.5}$$

The fact that $a_{ii}\leq 0$ is allowed shows that imaginary simple roots ($r^2 \leq 0$) appear in contrast to Kac-Moody algebras. There is again a denominator formula for GKM algebras which reads

$$e^{\rho}\prod_{r>0}(1-e^{r})^{\mathrm{mult}(r)}=\sum_{\sigma\in W}(\mathrm{sgn}(\sigma))\sigma(e^{\rho}\sum_{r}\epsilon(r)e^{r}) \tag{7.6}$$

where the correction factor on the right involves $\epsilon(r)$ which is $(-1)^n$ if r is the sum of n distinct pairwise orthogonal imaginary roots and zero otherwise.

As with hyperbolic Kac-Moody algebras, these formal definitions are of little practical utility without some method of determining the root multiplicities. For a few special GKM algebras it is possible in some cases to determine root multiplicities through the use of product formulae.

A more useful construction of a GKM algebra arises as the vertex operator algebra of physical string states in compactifications of string theory based on Lorentzian lattices, for example in toroidal compactifications of all string coordinates including time [20,77,23,25]. In our application the metric is the Narain metric rather than the spacetime metric but similar mathematical considerations apply.

We now discuss two specific examples related to the product representation of threshold corrections obtained in the previous section.

7.1. $\Pi^{1,1}$

We will first consider a GKM algebra with root lattice $\Pi^{1,1}$. Lattice vectors are labelled by a pair of integers (m,n) and the inner product of two lattice vectors is $(m,n)\cdot(m',n') = -(m'n+n'm)$. There is a single real simple root of length squared two which can be chosen to be $r_{-1}=(1,-1)$.

The Weyl group is generated by reflections in the real simple roots and so here is just the Z_2 transformation acting as

$$\sigma_{-1}(r) = r - (r\cdot r_{-1})r_{-1} \tag{7.7}$$

and thus takes $(m,n)\to(n,m)$.

A lattice Weyl vector is defined to be a lattice vector ρ which satisfies [78]

$$\rho\cdot r = -r^2/2 \tag{7.8}$$

for any simple real root. Applying this to $r=(1,-1)$ shows that lattice Weyl vectors must have the form $\rho=(m,m+1)$ for some integer m.

We now compare this to the Weyl vector ρ appearing in (4.21)

$$\rho = c(-1)(0,1) - \frac{c(0)}{24}(1,1) \tag{7.9}$$

where by the conditions given $F(q)=\sum_{n=-1}^{\infty} c(n)q^n$ is a modular function (that is weight zero). Thus $F(q)=(j(q)-744)+c(0)$. Taking $c(-1)=1$ and using in addition the condition that $24|c(0)$ we see that (7.9) gives a lattice Weyl vector.

Borcherds has investigated two different GKM algebras related to the $\Pi^{1,1}$ lattice. The first is the "fake" monster Lie algebra based on the Lie algebra of physical states on the lattice $\Pi^{25,1} = \Lambda_{\rm Leech} \oplus \Pi^{1,1}$. This algebra is just the algebra of physical string states in a covariant theory with the transverse degrees of freedom of the string compactified on the Leech lattice and the remaining space and time dimensions on the $\Pi^{1,1}$ lattice. The second theory is the monster Lie algebra based roughly speaking on the Lie algebra of physical states associated with the tensor product of the the $\Pi^{1,1}$ system with the monster vertex algebra of FLM based on a Z_2 orbifold of the Leech lattice [79]. In this case $c(0) = 0$ since all the massless transverse states are projected out and the Weyl vector is $\rho = (0, 1)$. In (4.20) we have a product over roots $r = (-\ell, -k)$ satisfying the conditions $k > 0$, $\ell \in Z$ or $k = 0$, $\ell > 0$. Since $c(0) = 0$ the only roots of the algebra (as compared to the lattice) appearing in this product are those with $k > 0$, $\ell \in Z$ (and with $\ell \geq -1$).

We will now show that the condition $r > 0$ encountered in the threshold corrections is the positive root condition for a GKM algebra with Weyl vector $\rho = (0, 1)$. According to a theorem in [71] if r is a positive root then $r^2 \leq -2\rho \cdot r$ with equality if and only if r is simple. With $r = (-\ell, -k)$ and $\rho = (0, 1)$ as above this gives the inequality

$$-2\ell k \leq -2\ell \tag{7.10}$$

which is satisfied for the roots $r = (-\ell, -k)$ with $k > 0$, $\ell \in Z$ and furthermore this shows that the simple roots are the real root $(1, -1)$ and the imaginary roots $(-\ell, -1)$ for $\ell > 0$.

Using the facts that the unique real simple root has multiplicity one and that there are no pairwise orthogonal simple roots the denominator formula (7.6) gives

$$p^{-1} \prod_{\ell \in Z, k > 0} (1 - p^k q^\ell)^{\mathrm{mult}(-\ell, -k)} = \sum_{\ell \geq -1, n \neq 0} \mathrm{mult}(-\ell, -1)(p^\ell - q^\ell) \tag{7.11}$$

with $p = e^{-(0,1)}$ and $q = e^{-(1,0)}$. As was shown by Borcherds [35] the undetermined root multiplicities can be determined indirectly based on the denominator formula (1.2). For Borcherds this root multiplicity arises in the construction of the monster Lie algebra. This leads to root multiplicities of $\mathrm{mult}(-\ell, -k) = c(k\ell)$ in (7.11) where the coefficients $c(n)$ appear in the expansion of the modular j function with constant term set equal to zero, $j(q) - 744 = \sum c(n) q^n$. Here we obtain similar products but the root multiplicity is now interpreted as the (Z_2 graded) multiplicity of BPS states. The formula (7.11) is precisely that appearing in (4.20) up to the $c_3(0)$ term after substituting the characters $p = e^{2\pi i T}$ and $q = e^{2\pi i U}$.

For the compactification constructed previously we have $c(0) = -984$ which leads to the Weyl vector $\rho = (41, 42)$. This does not seem to lead to a simple positive root condition for a GKM algebra. However the $\Pi^{1,1}$ case is rather exceptional in that the coefficient $c(0)$ is not determined by the conditions on $F(q)$. There is therefore some ambiguity in the precise interpretation of the product formula we have obtained related to the treatment of the $c(0)$ terms. If we explicitly evaluate these terms in the product then the remaining product is as described above a product over the positive roots of the monster Lie algebra. This ambiguity does not arise for lattices $\Pi^{s+1,1}$ with $s > 0$. We now turn to such an example.

7.2. $\Pi^{9,1}$

In this subsection we show that the condition $r > 0$ appearing in the threshold corrections is identical to the positive root condition for E_{10}. Recall that the condition $r > 0$ for $r = (\vec{b};, -\ell, -k)$ with $s = 8$ means:

$$\begin{aligned} k\ell - \tfrac{1}{2}\vec{b}^2 \geq -1 \quad & \text{and,} && (7.12a)\\ k > 0 \quad & \text{or,} && (7.12b)\\ k = 0, \quad \ell > 0 \quad & \text{or,} && (7.12c)\\ k = \ell = 0, \quad \vec{b} > 0 \quad & . && (7.12d) \end{aligned}$$

where we have taken into account that the multiplicities $c(-r^2/2)$ vanish for $r^2 > 2$ in (7.12*a*).

On the other hand, a choice of simple roots for the E_{10} lattice is given by taking:

$$r_i = (\vec{b}; 0, 0) = (\vec{\alpha}^{(i)}; 0, 0) \qquad i = 1, 8 \tag{7.13}$$

where $\vec{\alpha}^{(i)}$ is a set of simple roots of E_8, together with the extra root of E_9,

$$r_0 = (-\vec{\theta}; -1, 0), \tag{7.14}$$

($\vec{\theta}$ is the highest root of E_8), and

$$r_{-1} = (\vec{0}; 1, -1). \tag{7.15}$$

Given a set of simple roots r_μ the positive roots of E_{10} are simply

$$r = a_{-1}r_{-1} + a_0 r_0 + a_i r_i \tag{7.16}$$

with positive a_μ. We claim this set coincides with the set $r > 0$ above.

Several cases need to be checked. The only one that causes any difficulty is the case $k > 0, \ell > 0$. To check this suppose

$$k\ell \geq \tfrac{1}{2}\vec{b}^2 - 1. \tag{7.17}$$

Let

$$a_i \vec{\alpha}_i = \vec{b} + (k+\ell)\vec{\theta} \tag{7.18}$$

We then want to show that $a_i > 0$.

Proof: Let λ_i be the fundamental weights dual to α_i. Then

$$a_i = \vec{\lambda}_i \cdot \vec{b} + (k+\ell)\vec{\lambda}_i \cdot \vec{\theta}. \tag{7.19}$$

Now, by the Schwarz inequality and (7.17) we have

$$|\vec{\lambda}_i \cdot \vec{b}| \leq \sqrt{2k\ell + 2}|\lambda_i| \tag{7.20}$$

so

$$a_i \geq -\sqrt{2k\ell + 2}|\lambda_i| + (k+\ell)\vec{\lambda}_i \cdot \vec{\theta} \tag{7.21}$$

Now $\vec{\lambda}_i \cdot \vec{\theta} > 0$, and for $k > 0, \ell > 0$ we have:

$$(k+\ell)\vec{\lambda}_i \cdot \vec{\theta} \geq \sqrt{2k\ell + 2}|\lambda_i| \tag{7.22}$$

if we satisfy

$$(\vec{\lambda}_i \cdot \vec{\theta})^2 \geq \vec{\lambda}_i^2 \tag{7.23}$$

which is equivalent to

$$n_i^2 \geq G_{ii} \tag{7.24}$$

with G_{ii} the diagonal elements of the quadratic form of E_8 and n_i the numerical marks (or Coxeter labels) of E_8. This is true by inspection.

Moreover, the Weyl vector $\rho = (-\vec{\rho}, -31, -30)$ can be identified with the E_{10} lattice Weyl vector. To show this it suffices to check

$$\rho \cdot r_\mu = -\tfrac{1}{2}r_\mu^2 \qquad \mu = -1, 0, 1, \ldots 8 \tag{7.25}$$

As a nontrivial check on the above statements one can prove that the vectors which satisfy the inequality:

$$(r,r) \leq -2(\rho,r) \tag{7.26}$$

for all positive roots of a GKM algebra mentioned above are precisely the vectors satisfying (7.12a,b,c,d) above.

We thus conclude that the product formula we obtain through fundamental domain integrals which in turn determine the threshold corrections can be written in terms of a product over the positive roots of E_{10}. Furthermore, the product formula determines a (graded) multiplicity associated to the roots of E_{10}. The precise relation between these multiplicities and the root multiplicities of the E_{10} hyperbolic Kac-Moody algebra is not evident to us.

8. Some Limiting Cases

Further insight into the meaning of the formula for the prepotential can be gleaned from examining some limiting cases.

8.1. $M_{\rm string} \to \infty$

In this limit we expect to recover the standard formulae of global N=2 supersymmetry. On the moduli space $\mathcal{N}^{10,2}$ we consider a generic point $y_0 = (\vec{0}; y_+, y_-)$ on the embedded $\mathcal{N}^{2,2}$ submanifold. In the limit $M_{\rm string} \to \infty$ a fiber of the normal bundle of this submanifold is identified with the vectormultiplet moduli space of the global E_8 theory. Working near $\mathcal{H}^{1,1} \subset \mathcal{H}^{9,1}$ we let

$$y = y_0 + (\delta\vec{y}; 0, 0) \tag{8.1}$$

and in terms of dimensionful fields we let

$$\delta\vec{y} = \frac{\vec{a}}{M_{\rm string}} \tag{8.2}$$

In the global limit the constraint $2y_{+,2}y_{-,2} > (\delta\vec{y}_2)^2$ is trivially satisfied for fixed $\vec{a}$ and hence, after modding out by the duality group we have a coordinate $[\vec{a}] \in \underline{\mathbf{t}}(E_8) \otimes \mathbb{C}/W$, as expected.

Restoring string units, using (4.25), and the $N=2$ nonrenormalization theorems, the prepotential to all orders of perturbation theory is given by:

$$\mathcal{F} = M_{\rm string}^2\Big[-Sy^2 - \frac{1}{(2\pi)^4}\sum_{r>0} c_1(-r^2/2) Li_3(e^{2\pi i r\cdot y}) + \mathcal{A}(y)\Big] \tag{8.3}$$

where $\mathcal{A}(y)$ is a cubic polynomial. Now, in taking the $M_{\rm string}\to\infty$ limit (with $\vec{a}$ held fixed) we may extract the global prepotential from

$$\mathcal{F} \to M_{\rm string}^2 S y_0^2 + i\mathcal{F}^{\rm global} + \mathcal{O}(\frac{|\vec{a}|}{M_{\rm string}}) \tag{8.4}$$

In order to take the limit we note that the positive roots $r=(\vec{r};-\ell,-k)>0$ with $k\neq 0$ or $\ell\neq 0$ come in pairs with $\pm\vec{r}$. Using this and the limiting formula

$$Li_3(1-x) \to -\tfrac{1}{2}x^2\log x + \mathcal{O}(x^3\log x) \tag{8.5}$$

for the $k=\ell=0$ roots we obtain a sum only over the positive roots of E_8:

$$i\mathcal{F}^{\rm global} = \tilde{S}\vec{a}^2 - \frac{1}{8\pi^2}\sum_{\vec{\alpha}>0}(\vec{\alpha}\cdot\vec{a})^2\log\big[\frac{2\pi i\vec{\alpha}\cdot\vec{a}}{M_{\rm string}}\big] \tag{8.6}$$

in agreement with the standard one-loop prepotential of the global theory, where, for any group G we have the one-loop expressions [7,8]

$$\begin{aligned}\mathcal{F} &= \frac{i}{4\pi}\sum_{\vec{\alpha}>0}(\alpha\cdot A)^2\log\frac{(\alpha\cdot A)^2}{\Lambda^2}\\ K &= 2\Re\Big[i\mathcal{F}_i\overline{A^i}\Big]\\ &= -\frac{1}{\pi}\Re\Big[\sum_{\vec{\alpha}>0}|\alpha\cdot A|^2(\log\frac{(\alpha\cdot A)^2}{\Lambda^2}+1)\Big]\end{aligned} \tag{8.7}$$

In (8.6) the renormalized "classical" coupling is

$$\begin{aligned}\tilde{S} =& S + \frac{1}{8\pi^2}\sum_{r>0}c_1(-r^2/2)d(\vec{r}^2)\\ &\big[e^{2\pi i r\cdot y_0}Li_2(e^{2\pi i r\cdot y_0}) + e^{4\pi i r\cdot y_0}Li_1(e^{2\pi i r\cdot y_0})\big]\end{aligned} \tag{8.8}$$

the quantity $d(n)$ is defined by

$$\sum_{\vec{r}^2=n}(\vec{r}\cdot\vec{a})^2 = d(n)\vec{a}^2. \tag{8.9}$$

Finally, the cubic terms in $\mathcal{A}(y)$ vanish. The ambiguous quadratic terms with imaginary coefficients correspond to a change in scale of $M_{\rm string}$.

8.2. $M_{\rm string} \to 0$

One could also try taking the opposite, $M_{\rm string} \to 0$ limit. This raises many new conceptual questions [15] but, through the analogy with spontaneously broken gauge theory, might be expected to reveal fundamental underlying symmetries in string theory [80,23,24].

We must now introduce normalized fields for the other two moduli

$$y_\pm = \frac{a_\pm}{M_{\rm string}} \tag{8.10}$$

In the low energy theory the fields $a_\pm$ may be viewed as Higgs fields in the Cartan subalgebra of the enhanced $SU(2)\times SU(2)$ or $SU(3)$ gauge symmetry. Thus $M_{\rm string} \to 0$ holding a fixed is equivalent to $y \to \infty$. More fundamentally $a_\pm$ may be viewed as geometrical data on T^2. Consider the two-torus with vanishing B field and a diagonal metric $G_{11} = R_1^2$ and $G_{22} = R_2^2$. Then $T = 2iR_1R_2$ and $U = iR_2/R_1$. Thus the $y \to \infty$ limit is equivalent to a decompactification to *five* dimensions. In this limit the dominant terms in $\mathcal{F}$ are the cubic terms in $\mathcal{A}(y)$ [16].

By the automorphic properties of $\mathcal{F}$ the limit $y \to \infty$ is equivalent to a limit with $y \to 0$. Taking a formal $y \to 0$ limit leads us to the following suggestive formulae. We formally take the $y \to 0$ limit term-by-term using (8.5) and

$$\mathcal{P}(x) \to \pi x\bar{x}\log x + \mathcal{O}(x^3\log x). \tag{8.11}$$

Applying this to the formulae for the prepotential (8.3) we get:

$$f^{\rm perturbative} \to -S(a)^2 - \frac{1}{8\pi^2}\sum_{r>0} c(-r^2/2)(r\cdot a)^2 \log[\frac{2\pi i r\cdot a}{M_{\rm string}}] \tag{8.12}$$

and the Kahler potential becomes, in the $y \to 0$ limit:

$$K \to -\log\left[(\Im S)(-\Im a)^2 - \frac{1}{16\pi^2}\Re\left[\sum_{r>0} c(-r^2/2)|r\cdot a|^2 \log[\frac{2\pi i r\cdot a}{M_{\rm string}}]^2\right]\right] \tag{8.13}$$

If we furthermore take a special weak-coupling limit with $(\Im S)(-\Im a)^2 \to \infty$ then we recover the form of the global answer (8.7) but with the sum running over the positive roots of E_{10}.

---

[15] and perhaps should not be attempted without the full nonperturbative answer

[16] This cubic polynomial is probably related to the very special geometry of $D = 5, N = 2$ supergravity [81]

The limits (8.12)(8.13) are formal since the sums do not converge. Nevertheless comparison with the formulae for the global case suggests that the weak coupling 5-dimensional theory may be viewed, from the four-dimensional theory as an $N=2$ gauge theory for a generalized Kac-Moody algebra - most of which is spontaneously broken- with the full, infinite, spectrum of BPS states playing the role of a tower of gauge bosons. From the above formulae we see that the true root multiplicities are given by the coefficients $c(k\ell-\frac{1}{2}\vec{b}^2)$.

9. An algebra associated to BPS states

The results of sections 4-8 suggest that there is a GKM Lie algebra associated to the BPS states in D=4,N=2 heterotic compactifications. In this section we will construct a GKM Lie super-algebra with even elements associated to BPS vectormultiplets and the odd elements associated to BPS hypermultiplets. For the specific compactifications discussed previously this algebra has root lattice $\Pi^{1,1}$ or $\Pi^{9,1}$ and has root multiplicities related to those of the BPS states.

There are two obvious puzzles in trying to construct an algebra of BPS states. First, if we try to define an algebra by looking at the OPE of BPS vertex operators then it is clear that in general the OPE will contain operators for non-BPS states. In particular we will find operators with right-moving conformal dimension $\bar{h}>1/2$. Second, the physical BPS states are connected to a lattice of the form $\Pi^{s+2,2}$ while the products involved in threshold corrections involve only the lattice $\Pi^{s+1,1}$. We will see that the resolution to the first puzzle also provides the resolution to the second.

Consider the first point. In the Neveu-Schwarz sector a vertex operator which creates a BPS state is of the form:

$$V(z,\bar{z})=e^{ip_L X_L(z)+ip_R\tilde{X}_R(\bar{z})}\Phi_i(z,\bar{z}) \tag{9.1}$$

where X_R is the bosonic component of the free superfield in the internal $c=3$, $N=2$ theory, Φ is an operator of conformal dimensions $(h,\bar{h}=1/2)$ and

$$\tfrac{1}{2}p_L^2+h-1=\tfrac{1}{2}p_R^2 \qquad . \tag{9.2}$$

Note that Φ is chiral or anti-chiral primary with respect to the the total right-moving $N=2$ SCA with $U(1)$ charge ± 1. [17]

Using the decomposition (3.10) we can assign to each BPS vectormultiplet two distinguished vertex operators whose right-moving chiral primary field in the full internal $N=2$ theory is just the unit operator. In the lightcone gauge this vertex operator looks like:

$$V^A_{vm}(z) = e^{ip_L X_L(z) + ip_R \tilde{X}_R(\bar{z})} \Phi^A(z) \tilde{\psi}^\mu(\bar{z}) \tag{9.3}$$

where the index A runs over an infinite range and $\tilde{\psi}^\mu(\bar{z})$ is the right-moving transverse spacetime fermion field, i.e., $\mu = 2,3$. Now in the OPE of two such operators with opposite p_R we will find terms of the form $\prod \partial^n_{\bar{z}} X_R$. These operators create states with right-moving oscillators in the $c=3$ theory excited, that is states which are not BPS saturated. At this point we exploit the fact that the $c=3$ system is a free system. In (9.3) we can unambiguously drop the right-moving spacetime part of the vertex operator and change the right-moving free fields $X_R(\bar{z})$ to left-moving free fields which we will denote by $X_R(z)$ while at the same time retaining the Narain metric on (p_L, p_R). Technically, we have defined a mapping from the BPS states in the original light-cone gauge CFT of the heterotic string: [18]

$$\mathcal{C}^{2,0}_{\text{transverse}} \otimes \mathcal{C}^{0,3}_{\text{transverse}} \otimes \mathcal{C}^{0,3}_{\text{free N=(0,2)}} \otimes \mathcal{C}^{22,6}_{\text{N=(0,4)}} \tag{9.4}$$

to new leftmoving vertex operators of the form

$$V^A_{vm}(z,\bar{z}) \to \hat{V}^A_{vm}(z) \equiv e^{ip_L X_L(z) + ip_R X_R(z)} \Phi^A(z) \tag{9.5}$$

which live in the CFT:

$$\mathcal{C}^{2,0}_{\text{transverse}} \otimes \mathcal{C}^{22,6}_{\text{N=(0,4)}} \otimes \mathcal{C}^{2,0}_{\text{gaussian}} \tag{9.6}$$

Note that the physical state condition means that the left-moving conformal dimension of these vertex operators is

$$h + \tfrac{1}{2} p_L^2 - \tfrac{1}{2} p_R^2 = 1 \tag{9.7}$$

---

[17] In the examples we consider explicitly X_L is either the left-moving partner of X_R and $(p_L, p_R) \in \Pi^{2,2}$ or we extend X_L to include the lattice of the unbroken E_8 and $(p_L, p_R) \in \Pi^{10,2}$. In fact our considerations are more general and apply to any theory which has dimension one currents on the left, whether or not they arise from a free field construction, but for simplicity we will restrict ourselves to that case in what follows. We will also ignore cocycle factors in the vertex operators which are necessary to obtain a consistent operator algebra.

[18] Superscripts denote left and right conformal charges.

and hence these operators generate a current algebra based on the lattice $\Pi^{s+2,2}$. In terms of CFT correlators the structure constants for three vectormultiplets are

$$\left\langle \hat{V}^1_{vm}(1)\hat{V}^2_{vm}(2)\hat{V}^3_{vm}(3) \right\rangle \qquad . \tag{9.8}$$

We have now introduced two negative signature fields into our system and the techniques of BRST cohomology are the most appropriate to use when defining an algebra. We also need a mechanism to kill two sets of oscillator states and at the same time reduce the the $\Pi^{s+2,2}$ lattice to a $\Pi^{s+1,1}$ lattice. Such a mechanism is provided by string theory with local $N_{\rm ws}=2$ world-sheet supersymmetry, particularly in the heterotic version [82]. In heterotic $N_{\rm ws}=2$ string theory the leftmoving gauge algebra is a product $Vir \ltimes U(1)$. The gauged $U(1)$ current is of the form $J = v\cdot\partial X$ with $v \in \Pi^{s+2,2}$ where v is null. (For $\Pi^{2,2}$ and $\Pi^{10,2}$ there is a unique such null vector up to lattice automorphism.)

In the present context we do not have local $N_{\rm ws}=2$ world-sheet supersymmetry but we may nonetheless borrow these ideas to define a Lie algebra based on a $\Pi^{s+1,1}$ lattice which acts in a positive definite Hilbert space. We let L_n be the left-moving Virasoro operators constructed from the $c=(26,6)$ CFT defined by (9.6) and J_n the modes of the current defined in the previous paragraph. We then define the physical subspace of states as those annihilated by the L_n and J_n for $n\geq 0$. The J_0 constraint implies that physical states lie on a $\Pi^{s+1,1}$ sublattice of the $\Pi^{s+2,2}$ lattice parametrizing BPS states and the L_n and J_n constraints act to remove the negative norm states associated with the two timelike oscillators.

More formally, we enlarge the CFT to the BRST complex

$$\mathcal{C}^{2,0}_{\rm transverse}\otimes\mathcal{C}^{22,6}_{\rm N=(0,4)}\otimes\mathcal{C}^{2,0}_{\rm gaussian}\otimes\mathcal{C}^{-2,0}_{\xi\eta}\otimes\mathcal{C}^{-26,0}_{\rm b,c}\otimes\mathcal{C}^{2,0}_{x} \tag{9.9}$$

where the (ξ,η) system of $c=-2$ are the $U(1)$ ghosts. The last factor is a $c=2$ pair of free bosonic fields which is needed to ensure we are in the critical dimension. Now we may take the cohomology with respect to Q_{vir} and $Q_{U(1)}$ which are defined by:

$$\begin{aligned} Q_{vir} &= \oint\left[c\Big(T_L(z)+\tfrac{1}{2}(\partial x)^2-\tfrac{1}{2}(\partial X_R)^2+\eta\partial\xi\Big)-c\partial cb\right] \\ Q_{U(1)} &= \oint \xi v\cdot\partial X = \sum \xi_n v\cdot\alpha_{-n} \end{aligned} \tag{9.10}$$

where $T_L(z)$ is the leftmoving stress tensor of the $\mathcal{C}^{2,0}_{\rm transverse}\otimes\mathcal{C}^{22,6}_{\rm N=(0,4)}$ system. It is important to work in the "little Hilbert space" of the complex [83]. Thus we should also take cohomology with respect to $\oint \eta$. One important effect of this cohomology is on the allowed momentum. Since

$$\{\oint \eta, \oint \xi v \cdot \partial X\} = \oint v \cdot \partial X \qquad , \tag{9.11}$$

momenta must satisfy $p \cdot v = 0$ and are identified mod $p \sim p + \lambda v$. This removes momentum dependence in one $\Pi^{1,1}$ factor of the lattice. The cohomology with respect to the non-zero modes of J then kills two of the oscillators as desired. We can now define the vectormultiplet Lie algebra to be the algebra of physical states formed from the cohomology classes of vertex operators $\hat{V}^A_{vm}(z)$ with respect to the above cohomologies. By the no-ghost theorem the Virasoro cohomology has a positive definite contravariant form and thus the vectormultiplet Lie algebra is a generalized Kac-Moody Lie algebra (or Borcherds algebra) [35].

As we saw previously, threshold corrections depend on the difference between vectormultiplets and hypermultiplets. This suggests that we can extend this algebra to a Z_2 graded superalgebra with even elements corresponding to BPS vectormultiplets under the map (9.5) and the odd elements corresponding to hypermultiplets. Threshold corrections would then be interpreted in terms of a supertrace over the states of this superalgebra.

This extension can be accomplished as follows. From (3.10) we see that to each BPS hypermultiplet we can assign two distinguished vertex operators whose right-moving chiral primary field is a highest weight state for the $\mathcal{A}^{c=6}_{N=4}$ algebra and the unit operator for the $\mathcal{A}^{c=3}_{N=2}$ algebra.

$$V^{A,i}_{hm}(z,\bar{z}) = e^{ip_L X_L(z) + ip_R \tilde{X}_R(\bar{z})} \Phi^{A,i}(z,\bar{z}) \tag{9.12}$$

where A runs over an infinite range and i over a finite range (labelling the degeneracy of the $(\frac{1}{2},\frac{1}{2})$ representation of the $N=4, c=6$ theory.) We again construct a new vertex operator by making X_R left-moving and dropping the unit operator in the $c=3$ theory:

$$V^{A,i}_{hm}(z,\bar{z}) \to \hat{V}^{A,i}_{hm}(z,\bar{z}) \equiv e^{ip_L X_L(z) + ip_R X_R(z)} \Phi^{A,i}(z,\bar{z}) \tag{9.13}$$

This gives a vertex operator which is a holomorphic vertex operator multiplied by an antiholomorphic operator which is a primary chiral operator in the $\mathcal{A}^{c=6}_{N=4}$ algebra. To define a closed algebra we need to combine the left-moving current algebra with the graded commutative associative chiral algebra on the right. Since the $N=2$ chiral algebra splits according to (2.1), the bosonized $U(1)$ current also splits:

$$\sqrt{3} H^{tot} = H + \sqrt{2} H^{N=4} \tag{9.14}$$

so that the $N=4$ current is bosonized by: [19]

$$-2J_{+-} = -i\sqrt{2}\partial H^{N=4} \tag{9.15}$$

Since the $U(1)$ charge violation of the $N=4$ system is 2 the chiral ring of the $N=4$ theory gives us a map from $\hat{V}^{A,i}_{hm} \times \hat{V}^{B,j}_{hm} \to V^{C}_{vm}$, thus giving a $\mathbb{Z}_2$ grading of the desired type. Explicitly, the structure constants between two hypermultiplets and one vectormultiplet are:

$$\left\langle \hat{V}^1_{hm}(1)\hat{V}^2_{hm}(2)\hat{V}^3_{vm}(3)e^{-i\sqrt{2}H^{N=4}}(\bar{z}_4)\xi(z_0)\right\rangle \tag{9.16}$$

where z_0, z_4 can be inserted at any point $\neq z_1, z_2, z_3$.

More formally, we can take a twisted $N=4$ cohomology [84,85,86] on the right in addition to the Virasoro and $U(1)$ cohomologies on the left. From the (small) $N=4$ superconformal algebra, we have four odd supercurrents G_{AB}, $A, B = +, -$. Now we consider a twisting so that $T' = T - \partial J_{+-}$. Then G_{+A} has $h' = 1$ and G_{-A} has $h' = 2$. Thus, there are two BRST currents. In defining the topological states we should take *both* cohomologies

$$\tilde{Q}_A = \oint G_{+A} \tag{9.17}$$

The resulting cohomology states are chiral primary (not antichiral primary) with respect to all embedded N=2 algebras. On the resulting cohomology we can combine the leftmoving "Gerstenhaber bracket" with the rightmoving "Gerstenhaber product," [87][88], that is, we can take the product:

$$(V^1, V^2)(w,\bar{w}) \to \oint_w dz \lim_{\bar{z}\to\bar{w}} b_{-1}V^1(z,\bar{z}) \cdot V^2(w,\bar{w}) \tag{9.18}$$

Under this product the right-moving piece of two hypermultiplets maps to a right-moving chiral primary with charge +2. This does not correspond to a BPS vertex operator, but we can conjugate, or, equivalently, use the topological metric as in (9.16) above.

The vertex operators $\hat{V}$ associated to BPS states then form a graded Lie algebra of the generalized Kac-Moody type. If we introduce vectormultiplets with a multiplicity two then the root supermultiplicities of this algebra coincide with the degeneracies $c(-r^2/2)$ occurring in the product formula.

---

[19] In terms of the $N=2$ $U(1)$ current, $J^{(2)} = 2J_{+-}$.

10. Application to $N=2$ String Duality

There is now some evidence that $N=2$ heterotic string theories are dual to Type IIA or IIB string theory on special Calabi-Yau spaces [9,10]. These Calabi-Yau spaces have, roughly speaking, the structure of $K3$ surfaces fibered over rational curves [89][57]. The results presented here should allow for detailed tests of this duality in a broader class of models than has been considered so far. We hope to return to this issue in more detail elsewhere [90] but will make a few preliminary remarks here.

The heterotic theory we have considered arises from the symmetric embedding of the spin connection in the gauge group. In the Z_2 orbifold limit that we have used for explicit computations the massless spectrum consists of vector multiplets transforming in the adjoint representation of $E_8 \times E_7 \times SU(2) \times U(1)^4$ with a total of 388 states (including the graviphoton) and hypermultiplets transforming as

$$4(1,1,1)+8(1,56,1)+(1,56,2)+32(1,1,2). \tag{10.1}$$

with a total of 628 states. Note that $388-628=-240$ which matches the coefficient $c_1(0)$ in the $s=0$ case with unbroken E_8 gauge group. In order to compare with the spectrum of a possible dual Calabi-Yau space we can completely Higgs the $E_7 \times SU(2)$ gauge group (which does not change the difference between the number of vector and hypermultiplets) and break E_8 to $U(1)^8$ by turning on Wilson lines. This leaves us with gauge group $U(1)^{12}$ and 492 gauge neutral hypermultiplet fields [9,91]. The difference between the number of massless vector and hypermultiplets is then $12-492=-480$ which agrees with the coefficient $c_1(0)$ for $s=8$.

This massless spectrum would arise in Type II string theory on a Calabi-Yau space with $b_{11}=11$ and $b_{21}=491$ and hence with Euler number $\chi=-960$. Precisely such a Calabi-Yau family, denoted $X_{84}^{1,1,12,28,42}$ appears in the list of $K3$ fibrations given in [57]and is distinguished by having the maximal value of $|\chi|$ known for Calabi-Yau spaces.[20] These Calabi-Yau's are resolutions of degree 84 hypersurfaces in $\mathbb{P}_4^{1,1,12,28,42}$. A typical defining polynomial would be:

$$x_1^{84}+x_2^{84}+x_3^7+x_4^3+x_5^2=0. \tag{10.2}$$

Following the procedure in [89] we obtain a K3 fibration in the sense that there is a complete linear system $|L|$ whose divisors are K3 surfaces. In the present example we set

---

[20] We are grateful to D. Morrison for some very helpful remarks concerning these spaces.

$x_2 = \lambda x_1$ and define $y_1 = x_1^2$ to get a family $X_{42}^{1,6,14,21}$ of K3 surfaces realized as degree 42 polynomials in $\mathbb{P}_3^{1,6,14,21}$. For example (10.2) gives:

$$(1+\lambda^{84})y_1^{42} + x_3^7 + x_4^3 + x_5^2 = 0 \tag{10.3}$$

The K3 family $X_{42}^{1,6,14,21}$ possesses many beautiful and special properties, and has arisen before in the physics literature [92][93]. According to [93] the family is self-mirror with complexified Kahler cone $\cong \mathcal{H}^{9,1}$.

The formulae we have derived for the heterotic prepotential fit in well with the predictions of heterotic/type II duality. According to [9][57] the cohomology class in $H^2(X;Z)$ dual to a divisor L in the linear system defines the heterotic coordinate S. The remaining generators of the Kähler cone define coordinates $y \in \mathcal{H}^{9,1}$ and may be identified with the Kähler classes of $X_{42}^{1,6,14,21}$. The form of the S-dependence, Sy^2 of the classical prepotential follows from the structure of the $K3$-fibration, therefore, let us consider the third derivatives with respect to the y^A. The third derivative of the perturbative prepotential following from (4.25) is

$$\frac{\partial}{\partial y^A}\frac{\partial}{\partial y^B}\frac{\partial}{\partial y^C}\mathcal{F}^{\rm pert} = \frac{\tilde{d}^{10,2}_{ABC}}{64\pi^2} + \frac{1}{2\pi}\sum_{r>0} c_1(-r^2/2) r_A r_B r_C \frac{e^{-2\pi r\cdot y}}{1-e^{-2\pi r\cdot y}} \tag{10.4}$$

Note this is in precisely the right form expected for the counting of rational curves on a Calabi-Yau. Substituting the values of $\tilde{d}^{10,2}_{ABC}$ from the appendix we see that the third derivatives of $\mathcal{F}$ are indeed integers if we multiply $\mathcal{F}$ by $2\pi t$ where t is an integer. [21] Moreover, since $c_1(0) = \frac{1}{2}\chi$ we now recognize the constant term in (4.25):

$$-\frac{2\delta}{384\pi^2} = -\frac{c_1(0)\zeta(3)}{2(2\pi)^4}, \tag{10.5}$$

as the famous $\zeta(3)$ term appearing in the Calabi-Yau prepotential of [94] which is related to the four-loop renormalization of the Calabi-Yau sigma-model [95]. Comparing with the normalization of [94] we see that the properly normalized prepotential must be $4\pi\mathcal{F}$.

In view of this heterotic/type II duality makes some curious predictions for algebraic geometry, namely:

The cubic intersection product on $H^{1,1}$ for the family $X_{84}^{1,1,12,28,42}$ is governed by the structure of E_8 and may be computed from the expressions in $(A.51)$ and $(A.52)$

---

[21] In verifying this it is important to bear in mind the freedom to redefine $\mathcal{F}$ by (4.34). We hope to further investigate this chamber-dependence in the future [90].

by differentiation. Moreover, the rational curves in the family $X_{84}^{1,1,12,28,42}$ which are orthogonal to $c_1(L) \in H^2(X;\mathbb{Z})$ are parametrized by the positive roots $r > 0$ of E_{10} and appear with multiplicity $2c_1(-r^2/2)$ where $c_1(n)$ are the coefficients of the modular form $F(q) = E_6/\eta^{24}$.

If the rational curves are not isolated the integers must be interpreted as integrals over the moduli space, as is standard in topological field theory. These integers should be related to the numbers of rational curves on $K3$ surfaces in $X_{42}^{1,6,14,21}$.

Remarks entirely analogous to the above apply to the relation between the family of $K3$ fibrations $X_{24}^{1,1,2,8,12}$ discussed in [9][57] and the formula (4.36).

11. Concluding Speculations

We have shown that threshold corrections in $N = 2$ heterotic string theories are determined by the spectrum of BPS states and have provided evidence that there is a generalized Kac-Moody algebra associated with these states which governs the form of the threshold corrections. Various extensions of these results should prove very interesting.

It would be interesting to generalize the computations here to other backgrounds and more general dependence on the moduli. In particular, the computations done here involved the standard embedding of the spin connection on $K3$ in the gauge group. By repeating these calculations for different ranks of gauge group and for different topologies of gauge bundle it should be possible to recover interesting product formulae associated with a large number of hyperbolic and generalized Kac-Moody algebras. As examples, we expect to get an interesting algebra associated with $\Pi^{17,1}$ even with the symmetric embedding, and moreover we expect to recover the products occurring in the work of Feingold and Frenkel [96]. Indeed, the relevant products have already been suggested in [59].

Other embeddings and other choices of moduli dependence will also lead to modular integrals for congruence subgroups of $SL(2,\mathbb{Z})$. Given the presence (for the standard embedding) of the Monster Lie algebra and the j function, which is the Thompson series for the identity element of the Monster, it seems likely that these other embeddings will involve some of the other Thompson series for the Monster. It might be that the Monster provides a general classification of $N = 2$ heterotic string vacua.

The formula we have given for the perturbative heterotic prepotential fits in well with the proposed heterotic/type II string duality and therefore admits a natural extension to a full nonperturbative answer. In this paper we have shown that the special Kahler geometry of the $N = 2$ heterotic compactification is summarized by the prepotential:

$$\begin{aligned}\mathcal{F} = & - S(y, y) + \frac{1}{384\pi^2} i\tilde{d}^{s+2,2}_{ABC} y^A y^B y^C - \frac{\zeta(3)}{2(2\pi)^4} c_1(0) \\ & - \frac{1}{(2\pi)^4} \sum_{r>0} c_1(-r^2/2) Li_3(e^{2\pi i r\cdot y}) + \mathcal{F}^{\rm nonpert}(y, e^{2\pi i S})\end{aligned} \tag{11.1}$$

where $\mathcal{F}^{\rm nonpert}(y, e^{2\pi i S})$ has an analytic power series expansion in its second argument. As noted in the previous section, the formula for the prepotential as a sum of trilogarithms is *dictated* by the curve-counting formulae given a dual type II background. Thus, for backgrounds admitting a dual pair, the full nonperturbative prepotential will again be a sum of trilogarithms, with the replacement

$$e^{2\pi i r\cdot y} \rightarrow e^{2\pi i r\cdot y + 2\pi i n S}$$

in the argument of the trilogarithms. The sum will run over the full positive Kahler cone of the dual Calabi-Yau variety. In view of our results it is natural to speculate that this sum will again be a sum over positive roots of some interesting algebra.

The results of this paper should have two interesting mathematical applications in the context of heterotic/type II string duality. First, given an $N = 2$ dual string pair the results presented here combined with those of [97] suggest that special combinations of Ray-Singer torsions on dual Calabi-Yau spaces should admit infinite product representations. These considerations are undoubtedly related to the recent work of Jorgenson and Todorov [98]. Second, interesting recent work of Lian and Yau [99], has shown that certain mirror maps associated to one-parameter families of algebraic K3 surfaces are related to Thompson series. Since the heterotic string is a more natural home for the Monster it is tempting to speculate that the algebraic structure of BPS states together with string duality might provide a way to understand the observations of [99].

The algebraic structures we have begun to uncover should also shed new light on string duality. For example, it is tempting to speculate that different dual theories are simply different representations of an underlying algebraic structure much as in the different realizations of affine Kac-Moody algebras. It would be very interesting to try to identify generalized Kac-Moody algebras in Type II string theories on Calabi-Yau spaces arising as $K3$ fibrations. In addition to the applications to $N = 2$ string duality outlined in the previous section our work might also have relations to the conjectured duality between heterotic and Type I theories [14]. The exchange of $\bar{z}$ and z dependence needed to define the vertex algebra is reminiscent of world-sheet orbifolds considered in [100,101,102] and recently applied to Type I -heterotic duality in [103]. There is possibly an alternative point of view on our construction which might be interesting to pursue. In closed string theory there is a Virasoro algebra which generates transformations on the spatial string coordinate σ with generators $\hat{L}_n = L_n - \bar{L}_{-n}$. If we take the $\bar{L}_{-n}$ to be those of the free right-moving bosons and the L_n those of the left-moving degrees of freedom then the vertex operators we have defined are dimension one as a result of the condition (9.7). If we were then to complexify σ we would obtain a holomorphic vertex operator algebra. This suggests a complexification of the string world-sheet, something that has also been suggested in other contexts [104].

Finally, if the GKM algebra associated to BPS states that we have found is a gauge algebra, as it appears, then in analogy to the structure of $N = 2$ Yang-Mills theory the full nonperturbative prepotential should be governed by "monopoles" of this algebra and the monodromies given by an algebraic variety whose monodromy group is the Weyl group of this algebra. There is a natural candidate for these monopoles. Recall that the gauge structure became clearest in the $M_{\text{string}} \to 0$ limit, which coincides with a decompactification to 5 dimensions. We propose that the states being counted by the product representation of threshold corrections are five-branes wrapped around the remaining $K3 \times S^1$ internal space.

Acknowledgements

We thank the Aspen Center for Physics where this work was begun for providing a stimulating atmosphere. This work was supported in part by NSF Grant No. PHY 91-23780 and DOE grant DE-FG02-92ER40704.

We would like to thank S. Ferrara, A. Losev, J. Louis, D. Lüst, E. Martinec, D. Morrison, H. Nakajima, N. Nekrasov, H. Nicolai, M. O'Loughlin, S. Shatashvili, S. Shenker, A. Taormina, and A. Todorov for discussions. J.H owes special thanks to Neam Harvey.

Appendix A. Fundamental Domain Integrals

In this appendix we evaluate the following two integrals:

$$\mathcal{I}_{s+2,2}(y) \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2} \left[\left(\sum_{p \in \Gamma^{s+2,2}} q^{\frac{1}{2}p_L^2} \bar{q}^{\frac{1}{2}p_R^2} \right) F(q) - c(0) \right] \tag{A.1}$$

and

$$\tilde{\mathcal{I}}_{s+2,2}(y) \equiv \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2}\left[\Bigl(\sum_{p\in\Gamma^{s+2,2}} q^{\frac{1}{2}p_L^2}\bar{q}^{\frac{1}{2}p_R^2}\Bigr)F(q)(E_2-\frac{3}{\pi\tau_2})-\tilde{c}(0)\right] \tag{A.2}$$

These integrals are absolutely convergent for $y \in \mathcal{H}^{s+1,1}$ where y is not on the RQD's. They have logarithmic singularities on the RQD's and moreover, are $O(s+2,2;\mathbb{Z})$ invariant.

The notation is as follows: $\mathcal{F}$ is the fundamental domain for $SL(2,\mathbb{Z})$. $F(q)$ is a $SL(2,\mathbb{Z})$ modular function of weight $-s/2$ for $\mathcal{I}$ or of weight $-s/2-2$ for $\tilde{\mathcal{I}}$ which is holomorphic except at infinity where it has a first order pole. That is, we assume F has a Fourier expansion in terms of $q=e^{2\pi i\tau}$:

$$F(q)=\sum_{n=-1}^{\infty} c(n)q^n = c(-1)q^{-1}+c(0)+\cdots \tag{A.3}$$

and define $c(n)\equiv 0$, $n<-1$. We also define coefficients $\tilde{c}(n)$ through the expansion:

$$F(q)E_2(q)\equiv \sum_{n=-1}^{\infty}\tilde{c}(n)q^n \tag{A.4}$$

where $E_2(q)$ is the first of the series of Eisenstein functions:

$$E_2(q)=1-24\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n}=1-24\sum_{n=1}^{\infty}\sigma_1(n)q^n\ , \tag{A.5}$$

$$E_4(q)=1+240\sum_{n=1}^{\infty}\frac{n^3q^n}{1-q^n}=1+240\sum_{n=1}^{\infty}\sigma_3(n)q^n\ , \tag{A.6}$$

$$E_6(q)=1-504\sum_{n=1}^{\infty}\frac{n^5q^n}{1-q^n}=1-504\sum_{n=1}^{\infty}\sigma_5(n)q^n\ , \tag{A.7}$$

where

$$\sigma_k(n)=\sum_{d|n}d^k \tag{A.8}$$

is the sum of the kth powers of the divisors of n.

$E_2(\tau)$ is not modular covariant since it transforms with a shift. The combination

$$E_2(\tau)-\frac{3}{\pi\tau_2} \tag{A.9}$$

appearing in (A.2) on the other hand is not holomorphic but transforms covariantly.

The sums in (A.1) and (A.2)run over an even self-dual lattice $\Gamma^{s+2,2}(y) \cong \Pi^{s+2,2}$, and thus $s = 8t$ with t integer. As discussed in the text we write $\Pi^{8t+2,2} \cong \Pi^{8t,0} \oplus \Pi^{2,2}$ and write lattice vectors as

$$(\vec{b}; m_+, n_-; m_0, n_0) \tag{A.10}$$

with metric

$$(\vec{b}; m_+, n_-; m_0, n_0)^2 = \vec{b}^2 - 2m_+n_- + 2m_0n_0 \tag{A.11}$$

Also, for Narain compactifications we have

$$\begin{aligned} \tfrac{1}{2}(\vec{p}_L^2 - \vec{p}_R^2) &= \tfrac{1}{2}\vec{b}^2 - m_+n_- + m_0n_0 \\ \tfrac{1}{2}p_R^2 &= \frac{1}{-2(\Im y)^2}\left|\vec{b}\cdot\vec{y} - m_+y_- - n_-y_+ + m_0 - \tfrac{1}{2}n_0y^2\right|^2 \end{aligned} \tag{A.12}$$

Our conventions for the complex coordinates on the homogeneous space $SO(8t+2,2)/SO(8t+2)\times SO(2)$ are as in sec. 2.

A.1. Evaluation of $\mathcal{I}_{s+2,2}$

We will follow quite closely the calculation in the appendix of [42]. The general strategy for the evaluation of (A.1) and (A.2) is to perform first a Poisson resummation on the "momenta" m_+ and m_0 which leads to a sum over matrices A which are general two by two matrices with integer elements. The contributions of two matrices A, A' related by an element $V$of $SL(2,\mathbb{Z})$are related by modular transformation of τ by V which allows one to sum instead over orbits of $SL(2,\mathbb{Z})$and integrate over the images of the fundamental domain under the elements V that yield distinct matrices A when acting on a representative element of the orbit.

We now consider (A.1). Performing a Poisson resummation on m_+, m_0 gives

$$\tau_2 Z_\Gamma = \tau_2 \sum_{p\in\Gamma^{s+2,2}} q^{\frac{1}{2}p_L^2}\bar{q}^{\frac{1}{2}p_R^2} = \sum_{A\in Mat_{2\times 2}} T_\Gamma[A] \tag{A.13}$$

where

$$T_\Gamma[A] = \frac{-(\Im y)^2}{2y_{-,2}} \sum_{\vec{b}\in\Gamma^{8s,0}} q^{\frac{1}{2}\vec{b}^2} \exp\mathcal{G}, \tag{A.14}$$

$$\begin{aligned} \mathcal{G} = \Bigg[&\frac{\pi(\Im y)^2}{2y_{-,2}^2\tau_2}|\mathcal{A}|^2 - 2\pi i y_+ \det A + \frac{\pi}{y_{-,2}}\vec{b}\cdot(\vec{y}\tilde{\mathcal{A}} - \vec{y}^*\mathcal{A}) \\ &- \frac{\pi}{2y_{-,2}}n_0(\vec{y}^2\tilde{\mathcal{A}} - (\vec{y}^*)^2\mathcal{A}) + \frac{i\pi}{y_{-,2}^2}(\Im\vec{y})^2\ (n_- + n_0y_-^*)\mathcal{A}\Bigg] \end{aligned} \tag{A.15}$$

and

$$\begin{aligned}\mathcal{A} &= (\,1 \quad y_-\,)\begin{pmatrix} n_- & k_1 \\ n_0 & -k_2 \end{pmatrix}\begin{pmatrix} \tau \\ 1 \end{pmatrix} = (\,1 \quad y_-\,)\, A \begin{pmatrix} \tau \\ 1 \end{pmatrix} \\ \tilde{\mathcal{A}} &= (\,1 \quad y_-^*\,)\begin{pmatrix} n_- & k_1 \\ n_0 & -k_2 \end{pmatrix}\begin{pmatrix} \tau \\ 1 \end{pmatrix}\end{aligned} \tag{A.16}$$

Under modular transformations we have

$$\begin{aligned} A &\to A\begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \mathcal{A} &\to \frac{1}{c\tau+d}\mathcal{A}\end{aligned} \tag{A.17}$$

Unlike the simpler integral analyzed in [42], it is no longer obvious that the contribution from two matrices A related by the modular transformation (A.17) is given by the modular transformation $\tau' = \frac{a\tau+b}{c\tau+d}$. However for $s = 8t$ it is still true as can be seen by Poisson resummation on the $\Pi^{8t,0}$ lattice sum and using the fact that $\Pi^{8t,0}$ is even self-dual. Note that one could equally well have used a Poisson resummation on n_-, m_0 leading to the same expression with y_+ exchanged for y_-.

Following [42] we can split the sum on A into three orbits and correspondingly write the integrals as a sum of three integrals:

$$\begin{aligned}\mathcal{I}_{s+2,2} &= \mathcal{I}^0_{s+2,2} + \mathcal{I}^{nd}_{s+2,2} + \mathcal{I}^{dg}_{s+2,2} \\ \mathcal{I}^0_{s+2,2} &= \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} T_\Pi[A=0]F(q) \\ \mathcal{I}^{nd}_{s+2,2} &= 2\int_{\mathcal{H}} \frac{d^2\tau}{\tau_2^2} \sum_{0\le j<k, p\neq 0} T_\Pi[A = \begin{pmatrix} k & j \\ 0 & p \end{pmatrix}]F(q) \\ \mathcal{I}^{dg}_{s+2,2} &= \int_{\mathcal{S}} \frac{d^2\tau}{\tau_2^2}\left[\sum_{j,p}{}' T_\Pi[A = \begin{pmatrix} 0 & j \\ 0 & p \end{pmatrix}]F(q) - \tau_2 c(0)\theta(\tau\in\mathcal{F})\right]\end{aligned} \tag{A.18}$$

where $\mathcal{H}$ is the upper half plane and $\mathcal{S}$ is the strip $\{\tau \in \mathcal{H}, |\tau_1| < 1/2\}$. These three terms correspond to the zero orbit $A = 0$, the non-degenerate orbit with $detA \neq 0$ and the degenerate orbit with $detA = 0$ and a particular choice of representative matrices from each orbit.

In deriving (A.18) we have made an important exchange of summation and integration. Because of the tachyon divergences $\sim q^{-1}$ which exist before the $L_0 = \bar{L}_0$ projection this exchange can be invalid. Indeed, it is clear that if $y_{+,2} > 2y_{-,2}$ then one must use the expressions (A.13),(A.14)(A.15). On the other hand, if $y_{-,2} > 2y_{+,2}$, one must use the other Poisson summation with $y_+ \leftrightarrow y_-$. In fact, we know that the expressions must change as we cross the wall of the Weyl chamber $y_{-,2} = y_{+,2}$. [22]

The $A = 0$ orbit may be evaluated using [105] [23] .

$$\frac{-(\Im y)^2}{2y_{-,2}} \int_{\mathcal{F}} \frac{d^2\tau}{\tau_2^2} \vartheta_{\Pi^{8t,0}} F(q) = \frac{-(\Im y)^2}{2y_{-,2}} \frac{1}{\pi} [G_2(q)\vartheta_{\Pi^{8t,0}} F(q)]_{q^0} \tag{A.19}$$

For $t = 1$ which will be of most relevance for our discussion we have

$$\frac{1}{\pi}[G_2(q)\vartheta_{\Pi^{8,0}} F(q)]_{q^0} = \frac{\pi}{3}(c(0) + 216c(-1)) \tag{A.20}$$

The non-degenerate orbit is evaluated using the representative matrix

$$A_0 = \begin{pmatrix} k & j \\ 0 & p \end{pmatrix} \tag{A.21}$$

with $0 \le j < k$ and $p \neq 0$. The τ_1 integral is gaussian and the sum on j is then trivial. The integral over τ_2 gives a representation of the Bessel function $K_{1/2}$ which is elementary. The sum on p yields a logarithm so that the contribution from the non-degenerate orbit is given by

$$-2\log\left|\prod_{\vec{b}\in\Pi^{8t,0}}\ \prod_{k>0,\ell\in\mathbb{Z}}\left(1 - e^{2\pi i r\hat{\cdot}y}\right)^{c(k\ell - \frac{1}{2}\vec{b}^2)}\right|^2 \tag{A.22}$$

where we have introduced a “hatted dot product ” defined by

$$r\hat{\cdot}y \equiv \Re\left[(\vec{b}\cdot\vec{y} + \ell y_- + k y_+)\right] + i\left|\Im\left[(\vec{b}\cdot\vec{y} + \ell y_- + k y_+)\right]\right| \tag{A.23}$$

when $k > 0$, and, when $k = 0$:

$$\begin{aligned} r\hat{\cdot}y &\equiv r\cdot y - N(r,y)y_- \\ N(r,y) &= \operatorname{sgn}(\vec{b})\left[\operatorname{sgn}(\vec{b})\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}\right] \end{aligned} \tag{A.24}$$

where $[\cdot]$ is the greatest integer function.

To evaluate the contribution from the degenerate orbit we take

$$A_0 = \begin{pmatrix} 0 & j \\ 0 & p \end{pmatrix} \tag{A.25}$$

---

[22] Hence there is a puzzle about why the exchange of sum and integral is not valid for one Poisson summation in the range $2 > y_{+,2}/y_{-,2} > 1$. We have not resolved this point.

[23] After correcting a factor two error in eqn. 9.38 of [105].

with j, p not both zero. After evaluating the τ_1 integral the sum on j may be performed using a Sommerfeld-Watson transformation:

$$\begin{aligned}\sum_{j=-\infty}^{\infty} \frac{e^{i\theta j}}{(j+B)^2+C^2} &= \frac{\pi}{C} e^{-i\theta(B-iC)} \frac{1}{1-e^{-2\pi i(B-iC)}} \\ &\quad + \frac{\pi}{C} e^{-i\theta(B+iC)} \frac{e^{2\pi i(B+iC)}}{1-e^{2\pi i(B+iC)}} \\ &\quad C > 0, \qquad 0 \le \theta \le 2\pi\end{aligned} \tag{A.26}$$

A special case is

$$\sum_{j=1}^{\infty} \frac{\cos\theta j}{j^2} = \frac{\theta(\theta-2\pi)}{4} + \frac{\pi^2}{6} \tag{A.27}$$

for $0 \le \theta \le 2\pi$.

To write the final answer it is useful to introduce the polylogarithm functions:

$$\begin{aligned} Li_1(x) &= \sum_{j=1}^{\infty} \frac{x^j}{j} = -\log(1-x) \\ Li_2(x) &= \sum_{j=1}^{\infty} \frac{x^j}{j^2} = \int_0^1 \frac{dt}{t} \frac{1}{(1-xt)} \\ Li_3(x) &= \sum_{j=1}^{\infty} \frac{x^j}{j^3} = -\int_0^1 \frac{dt}{t} \int_0^1 \frac{ds}{s} \log(1-xts) \end{aligned} \tag{A.28}$$

The general answer for the integral (A.1) is then

$$\begin{aligned}\mathcal{I}_{8t+2,2}(y) &= \frac{-(\Im y)^2}{2y_{-,2}} \frac{\pi}{3} [E_2(q)\vartheta_{\Pi^{8t,0}} F(q)]_{q^0} \\ &\quad + c(0)\left(-\log[-(\Im y)^2] + \frac{\pi}{3} y_{-,2} - \mathcal{K}\right) \\ &\quad + c(-1) \frac{2y_{-,2}}{\pi} \sum_{\vec{b}^2=2} \Re\left[Li_2\big(e^{2\pi i \frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}}\big)\right] \\ &\quad + 4\Re\left\{\sum_{r>0}\left[c(kl-\tfrac{1}{2}\vec{b}^2) Li_1(e^{2\pi i r\hat{\cdot} y})\right]\right\}\end{aligned} \tag{A.29}$$

Here $r > 0$ means:

1. $k > 0$ or,
2. $k = 0$, $\ell > 0$ or,

3. $k = \ell = 0$, $\vec{b} > 0$.

The first line of (A.29) comes from $A = 0$. The second and third lines come from the degenerate orbit, and the last line comes from the nondegenerate and degenerate orbits. The constant $\mathcal{K}$ is

$$\mathcal{K} = \log[\frac{4\pi}{\sqrt{27}} e^{1-\gamma_E}] \tag{A.30}$$

where γ_E is the Euler-Mascheroni constant.

We now consider two special cases.

$\underline{s = 0}$: In this case it is customary to denote $y_+ = T, y_- = U$. The general formula (A.29) then gives

$$\begin{aligned}\mathcal{I}_{2,2}(T,U) &= c(0)[-\log[2T_2U_2] - \mathcal{K}] + c(0)\frac{\pi}{3}(T_2 + U_2) \\ &\quad - 2\log\left| e^{-2\pi i T c(-1)} \prod_{r>0} \left(1 - e^{2\pi i (kT+\ell U)}\right)^{c(k\ell)} \right|^2\end{aligned} \tag{A.31}$$

Here $r > 0$ means $k > 0, \ell \in \mathbb{Z}$, or $k = 0, \ell > 0$. We have also assumed that $T_2 > U_2$. In this case one may replace $r \hat{\cdot} y$ with $r \cdot y$ since they only differ for $k = 1, \ell = -1$ and in this case $r \hat{\cdot} y = (T_1 - U_1) + i|T_2 - U_2|$ which equals $r \cdot y$ for $T_2 > U_2$. For $U_2 > T_2$ one should interchange U and T in (A.31).

The last two terms in (A.31) may be written as

$$-2\log|j(T) - j(U)|^2 - 2\log|\eta(T)\eta(U)|^{2c(0)} \tag{A.32}$$

using the product formula (1.2).

$\underline{s = 8}$: In this case we have, without loss of generality,

$$F(q) = \frac{E_4^2}{\eta^{24}} \tag{A.33}$$

The finite sums over simple roots of $\Pi^{8,0}$ can be simplified. In (A.29) we have a term

$$\Re\Big[Li_2(e^{i\theta})\Big] = \sum_{j=1}^{\infty} \frac{\cos\theta j}{j^2} = \frac{\theta(\theta - 2\pi)}{4} + \frac{\pi^2}{6} \qquad 0 \le \theta \le 2\pi \tag{A.34}$$

In order to apply this to the third line of (A.29) we need to be careful to take care of the range of the angle $\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}$. Since, by assumption, y is not at an enhanced symmetry point this quantity is not integral. If we further take y to be in the fundamental Weyl chamber (5.12) we get:

$$2c(-1)\frac{2y_{-,2}}{\pi} \sum_{\vec{b}^2=2,\vec{b}>0} \left[\frac{\pi^2}{3} - 2\pi^2\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}} + 2\pi^2(\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}})^2\right] \tag{A.35}$$

We can evaluate the sum over E_8 root vectors. The Weyl vector of E_8 is $\vec{\rho} \equiv \frac{1}{2}\sum_{\vec{b}^2=2,\vec{b}>0}\vec{b}$ and the quadratic sum can be evaluated [106] to give

$$\sum_{\vec{b}^2=2}(\vec{b}\cdot\vec{v})^2 = 60\vec{v}^2 \tag{A.36}$$

To summarize, under the assumption (5.12) we get the product:

$$\begin{aligned}
\mathcal{I}_{10,2}(y) &= -2\log\left|\Phi(y)\right|^2 + c(0)\left(-\log[-(\Im y)^2] - \mathcal{K}\right)\\
\Phi(y) &= e^{2\pi i\rho\cdot y}\prod_{r>0}\left(1-e^{2\pi i r\hat{\cdot} y}\right)^{c(k\ell-\frac{1}{2}\vec{b}^2)}\\
\rho &= -(\vec{\rho};31,30)
\end{aligned} \tag{A.37}$$

In general the integral $\mathcal{I}_{s+2,2}(y)$ has a Weyl vector. This means that the cubic terms cancel, as can be derived using theorems 6.2 and 10.3 of Borcherds [29]. Moreover, the Weyl vector ρ appearing in the product formulae of [29] can be extracted from the terms linear in y. Abstractly we have - for all $t \geq 1$:

$$y_{+,2}\frac{\pi}{3}[E_2\vartheta_{\Pi^{8t,0}}F(\tau)]_{q^0} + y_{-,2}\frac{\pi}{3}[\vartheta_{\Pi^{8t,0}}F(\tau)]_{q^0} - 4\pi\sum_{\vec{b}>0}c(-\vec{b}^2/2)\vec{b}\cdot\Im\vec{y}. \tag{A.38}$$

This agrees exactly with the expression in theorem 10.4 of [29].

A.2. Evaluation of $\tilde{\mathcal{I}}$

We now turn to an evaluation of (A.2). We again decompose the sum over A into a sum over orbits of $SL(2,\mathbb{Z})$. The $A=0$ orbit is done using [105]

$$\int_{\mathcal{F}}\frac{d^2\tau}{\tau_2^2}(\hat{G}_2(\tau))^k F(q) = \frac{1}{\pi(k+1)}\left[(G_2(\tau))^{k+1}F(q)\right]_{q^0} \tag{A.39}$$

The evaluation of the non-degenerate orbit proceeds as before except that the τ_2 integral now gives a representation of $K_{3/2}$ and in the sum over p one makes the replacement

$$\frac{1}{|p|} \to \frac{\left|\Im\left[(\vec{b}\cdot\vec{y}+\ell y_- + k y_+)\right]\right|}{p^2} + \frac{1}{2\pi|p|^3} \tag{A.40}$$

Thus, instead of logarithms we get $Li_2(x)$ and $Li_3(x)$. The evaluation of the degenerate orbit also proceeds as before except that the sum on j now involves

$$\sum_{j=-\infty}^{\infty} \frac{e^{i\theta j}}{((j+B)^2+C^2)^2} = -\frac{1}{2C}\frac{\partial}{\partial C}\sum_{j=-\infty}^{\infty} \frac{e^{i\theta j}}{((j+B)^2+C^2)} \tag{A.41}$$

which thus reduces to (A.26). The result is

$$\begin{aligned}
\tilde{\mathcal{I}}_{8t+2,2}(y) =& \frac{-(\Im y)^2}{2y_{-,2}}\frac{\pi}{6}[E_2^2(q)\vartheta_{\Pi^{8t,0}}F(q)]_{q^0} + \tilde{c}(0)(-\log[-(\Im y)^2] + \frac{\pi}{3}y_{-,2} - \mathcal{K}) \\
&+c(0)\frac{2\pi}{15}\frac{y_{-,2}^3}{(\Im y)^2} + \frac{6}{\pi^2}\frac{c(0)}{(\Im y)^2}\zeta(3) \\
&+\sum_{\vec{b}^2=2}\Re\left[\tilde{c}(-1)\frac{2y_{-,2}}{\pi}Li_2\big(e^{2\pi i\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}}\big) + c(-1)\frac{12y_{-,2}^3}{\pi^3(\Im y)^2}Li_4\big(e^{2\pi i\frac{\vec{b}\cdot\Im\vec{y}}{y_{-,2}}}\big)\right] \\
&+4\Re\left\{\sum_{r>0}\left[\tilde{c}(kl-\tfrac{1}{2}\vec{b}^2)Li_1(e^{2\pi i r\hat{\cdot} y}) + \frac{6}{\pi(\Im y)^2}c(kl-\tfrac{1}{2}\vec{b}^2)\mathcal{P}(r\hat{\cdot} y)\right]\right\}
\end{aligned} \tag{A.42}$$

where we have also introduced the function

$$\mathcal{P}(x) = \Im(x)Li_2(e^{2\pi i x}) + \frac{1}{2\pi}Li_3(e^{2\pi i x}). \tag{A.43}$$

In fact, this integral can be written in the form:

$$\begin{aligned}
\tilde{\mathcal{I}}_{8t+2,2}(y) =& 4\Re\left\{\sum_{r>0}\left[\tilde{c}(kl-\tfrac{1}{2}\vec{b}^2)Li_1(e^{2\pi i r\hat{\cdot} y}) + \frac{6}{\pi(\Im y)^2}c(kl-\tfrac{1}{2}\vec{b}^2)\mathcal{P}(r\hat{\cdot} y)\right]\right\} \\
&+\tilde{c}(0)\bigg(-\log[-y_2)^2] - \mathcal{K}\bigg) + \frac{1}{(y_2)^2}[\tilde{d}_{ABC}^{8t+2,2}y_2^A y_2^B y_2^C + \delta]
\end{aligned} \tag{A.44}$$

where the constant is

$$\delta = \frac{6}{\pi^2}c(0)\zeta(3) \tag{A.45}$$

and $\tilde{d}$ is a real symmetric tensor, which depends on the Weyl chamber.

We now consider the two special cases $s=0$ and $s=8$. For the case $s=0$ without loss of generality we have

$$F = \frac{E_4E_6}{\eta^{24}} \tag{A.46}$$

With $y_+ = T, y_- = U$ as before we have

$$\begin{aligned}\tilde{\mathcal{I}}_{2,2}(y) =&4\Re\Bigg\{\sum_{r>0}\bigg[\tilde{c}(kl)Li_1(e^{2\pi i r\hat{\cdot}y}) - \frac{3}{\pi T_2 U_2}c(kl)\mathcal{P}(r\,\hat{\cdot}\,y)\bigg]\Bigg\}\\ -\frac{\delta}{2T_2U_2}&-264[-\log[2y_{+,2}y_{-,2}]-\mathcal{K}]\\ &-48\pi T_2 - 88\pi U_2 + 16\pi\frac{U_2^2}{T_2} \qquad T_2 > U_2\\ &-48\pi U_2 - 88\pi T_2 + 16\pi\frac{T_2^2}{U_2} \qquad U_2 > T_2\end{aligned} \tag{A.47}$$

As discussed above $r\,\hat{\cdot}\,y$ is only different from $r\cdot y$ for $k=1, \ell=-1$ and in this case it is $(T_1 - U_1) + i|T_2 - U_2|$.

When $s = 8$, we have

$$F = \frac{E_6}{\eta^{24}} \tag{A.48}$$

To calculate $\tilde{\mathcal{I}}$ we need

$$\Re\Big[Li_4(e^{i\theta})\Big] = \sum_{j=1}^{\infty}\frac{\cos\theta j}{j^4} = \frac{\pi^4}{90} - \frac{1}{48}\theta^2(2\pi-\theta)^2 \tag{A.49}$$

valid for $0 \le \theta \le 2\pi$. We also need the E_8 root sum

$$\sum_{\vec{b}^2=2,\vec{b}>0}(\vec{b}\cdot\vec{v})^4 = 18(\vec{v}^2)^2 \tag{A.50}$$

where the quartic sum can be evaluated using formulae in [106]. Again specializing to the Weyl chamber (5.12) we can write the answer as in (A.44) with the symmetric tensor $\tilde{d}$ is determined by:

$$\begin{aligned}\tilde{d}^{10,2}_{ABC}y_2^A y_2^B y_2^C =&-8\pi\Big[(\vec{\rho}\cdot\vec{y}_2 + 41y_{-,2} + 42y_{+,2})(y_2)^2\\ -2\sum_{\vec{b}^2=2,\vec{b}>0}&(\vec{b}\cdot\vec{y}_2)^3 + 60y_{-,2}^2 y_{+,2} + 72y_{+,2}^2 y_{-,2} + 4(y_{-,2})^3\Big]\end{aligned} \tag{A.51}$$

Finally, we need an explicit formula for the cubic sum over E_8 roots. This is given by

$$\begin{aligned}\sum_{\vec{b}^2=2,\vec{b}>0}(\vec{b}\cdot\vec{v})^3 =&6\,v(1)^2\,v(2)+2\,v(2)^3+6\,v(1)^2\,v(3)+6\,v(2)^2\,v(3)\\&+4\,v(3)^3+6\,v(1)^2\,v(4)+6\,v(2)^2\,v(4)+6\,v(3)^2\,v(4)\\&+6\,v(4)^3+6\,v(1)^2\,v(5)+6\,v(2)^2\,v(5)+6\,v(3)^2\,v(5)\\&+6\,v(4)^2\,v(5)+8\,v(5)^3+6\,v(1)^2\,v(6)+6\,v(2)^2\,v(6)\\&+6\,v(3)^2\,v(6)+6\,v(4)^2\,v(6)+6\,v(5)^2\,v(6)+10\,v(6)^3\\&+6\,v(1)^2\,v(7)+6\,v(2)^2\,v(7)+6\,v(3)^2\,v(7)+6\,v(4)^2\,v(7)\\&+6\,v(5)^2\,v(7)+6\,v(6)^2\,v(7)+12\,v(7)^3+30\,v(1)^2\,v(8)\\&+30\,v(2)^2\,v(8)+30\,v(3)^2\,v(8)+30\,v(4)^2\,v(8)+30\,v(5)^2\,v(8)\\&+30\,v(6)^2\,v(8)+30\,v(7)^2\,v(8)+22\,v(8)^3\end{aligned}\tag{A.52}$$

where $v(i)$ is the i^{th} component of v in the usual Cartesian basis for E_8 roots.

References

[1] C. Montonen and D. Olive, “Magnetic monopoles as gauge particles? ”Phys. Lett. **72B** (1977) 117; P. Goddard, J. Nuyts and D. Olive, “Gauge theories and magnetic charge,” Nucl. Phys. **B125** (1977) 1.

[2] E. Witten and D. Olive, “Supersymmetry algebras that include topological charges,” Phys. Lett. **78B** (1978) 97.

[3] H. Osborn, “Topological charges for $N = 4$ supersymmetric gauge theories and monopoles of spin 1,” Phys. Lett. **83B** (1979) 321.

[4] A. Sen, “Dyon-monopole bound states, selfdual harmonic forms on the multi-monopole moduli space, and $SL(2, Z)$ invariance in string theory,” Phys. Lett. **329** (1994) 217, hep-th/9402032.

[5] N. Seiberg and E. Witten, “Electric-magnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang-Mills theory,” Nucl. Phys. **B426** (1994) 19; (E) **B340** (1994) 485, hep-th/9407087.

[6] N. Seiberg and E. Witten, “ Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD, ” Nucl. Phys. **B431** (1994) 484, hep-th/9408099.

[7] P. C. Argyres and A. E. Faraggi, “The vacuum structure and spectrum of $N = 2$ supersymmetric $SU(n)$ gauge theory, ” Phys. Rev. Lett. **74** (1995) 3931, hep-th/9411057.

[8] A. Klemm, W. Lerche and S. Theisen, “Nonperturbative effective actions of $N = 2$ supersymmetric gauge theories,” hepth-9505150.

[9] S. Kachru and C. Vafa, “Exact results for $N = 2$ compactifications of heterotic strings, ” Nucl. Phys. **B450** (1995) 69, hep-th/9505105.

[10] S. Ferrara, J. A. Harvey, A. Strominger, C. Vafa , “Second-Quantized Mirror Symmetry, ” hep-th/9505162.

[11] A. Dabholkar and J. A. Harvey, “Nonrenormalization of the superstring tension,” Phys. Rev. Lett. **63** (1989) 478; A. Dabholkar, G. Gibbons, J. A. Harvey and F. Ruiz Ruiz, “Superstrings and solitons,” Nucl. Phys. **B340** (1990) 33.

[12] J. Schwarz and A. Sen, “Duality symmetries of 4-D heterotic strings,” Phys. Lett. **B312** (1993) 105, hep-th/9305185.

[13] C. Hull and P. Townsend, “Unity of superstring dualities,” Nucl. Phys. **B438** (1995) 109, hep-th/9410167.

[14] E. Witten, “String theory dynamics in various dimensions,” Nucl. Phys. **B443** (1995) 85, hep-th/9503124.

[15] A. Strominger, “Massless black holes and conifolds in string theory,” Nucl. Phys. **B451** (1995) 96; hep-th/9504090.

[16] B. R. Greene, D. R. Morrison and A. Strominger, “Black hole condensation and the unification of string vacua,” Nucl. Phys. **B451** (1995) 109; hep-th/9504145

[17] J. Polchinski, “Dirichlet-Branes and Ramond-Ramond charges” hep-th/9510017.

[18] For reviews and further references see M. Duff, R. Khuri and J. X. Lu, "String Solitons" Phys. Rep. **259** (1995) 213, hep-th/9412184; A. Sen, "Strong-weak coupling duality in four-dimensional string theory," Int. J. Mod. Phys. **A9** (1994) 3707.

[19] I. Frenkel, "Representations of Kac-Moody algebras and dual resonance models," in *Applications of Group Theory in Physics and Mathematical Physics*, Vol. 21, Lectures in Applied Mathematics, M. Flato, P. Sally, G. Zuckerman, eds. AMS 1985.

[20] P. Goddard and D. Olive, "Algebras, Lattices, and Strings," in *Vertex operators in mathematics and physics*," ed. J. Lepowsky et. al. Springer-Verlag, 1985.

[21] E. Witten, "Topological tools in ten dimensional physics, " in *Unified String Theories*, eds. M. Green and D. Gross, World Scientific, Singapore, 1986.

[22] A. Giveon and M. Porrati, "Duality invariant string algebra and $D = 4$ effective actions, " Nucl. Phys. **B355** (1991) 422.

[23] G. Moore, "Finite in All Directions, " hep-th/9305139; G. Moore, "Symmetries and symmetry-breaking in string theory," hep-th/9308052.

[24] G. Moore, "Symmetries of the Bosonic String S-Matrix," hep-th/9310026; Addendum to: "Symmetries of the Bosonic String S-Matrix," hep-th/9404025.

[25] R.W. Gebert, "Introduction to Vertex Algebras, Borcherds Algebras, and the Monster Lie Algebra," Int. J. Mod. Phys. **A8** (1993) 5441, hep-th/9308151.

[26] R. W. Gebert and H. Nicolai, " On E_{10} and the DDF construction," hep-th/9406175.

[27] A. Giveon, M. Porrati and E. Rabinovici, "Target space duality in string theory," Phys. Rep. **244** (1994) 77;hep-th/9401139

[28] R.W. Gebert, H. Nicolai and P.C. West, "Multistring Vertices and Hyperbolic Kac Moody Algebras," hep-th/9505106.

[29] R. Borcherds, "Automorphic forms on $O_{s+2,2}(R)$ and infinite products," Invent. Math. **120**(1995) 161.

[30] R. Borcherds, "Automorphic forms on $O_{s+2,2}(R)^+$ and generalized Kac-Moody algebras," contribution to the Proceedings of the 1994 ICM, Zurich.

[31] V. A. Gritsenko, V. V. Nikulin, "Siegel automorphic form corrections of some Lorentzian Kac–Moody Lie algebras, " alg-geom/9504006.

[32] V. V. Nikulin, "Reflection groups in hyperbolic spaces and the denominator formula for Lorentzian Kac–Moody Lie algebras," alg-geom/9503003.

[33] V. A. Gritsenko, V. V. Nikulin, "K3 Surfaces, Lorentzian Kac-Moody Algebras, and Mirror Symmetry," alg-geom/9510008.

[34] R. E. Borcherds, "The monster Lie algebra," Adv. Math. **83** No. 1 (1990).

[35] R. Borcherds,"Monstrous moonshine and monstrous Lie superalgebras," Invent. Math. **109**(1992) 405.

[36] J. H. Conway and S. P. Norton, "Monstrous moonshine," Bull. London Math. Soc. **11** (1979) 308.

[37] G. L. Cardoso, D. Lüst and T. Mohaupt, “Threshold corrections and symmetry enhancement in string compactifications,” Nucl. Phys. **B450** (1995) 115, hep-th/9412209.

[38] B. de Wit, V. Kaplunovsky, J. Louis and D. Lüst, “Perturbative Couplings of Vector Multiplets in $N = 2$ Heterotic String Vacua,” Nucl. Phys. **B451** (1995) 53, hep-th/9504006.

[39] I. Antoniadis, S. Ferrara, E. Gava, K.S. Narain and T.R. Taylor, “Perturbative Prepotential and Monodromies in N=2 Heterotic Superstring,” Nucl. Phys. **B447** (1995) 35, hep-th/9504034.

[40] J. Minahan, Nucl. Phys. **B298** (1988) 36.

[41] V. Kaplunovsky, “One loop threshold effects in string unification,” Nucl. Phys. **B307** (1988) 145, revised in hep-th/9205070.

[42] L. Dixon, V. S. Kaplunovsky and J. Louis, “Moduli-dependence of string loop corrections to gauge coupling constants, ” Nucl. Phys. **B329** (1990) 27.

[43] I. Antoniadis, E. Gava, K.S. Narain, “Moduli corrections to gravitational couplings from string loops,” Phys. Lett. **B283** (1992) 209, hep-th/9203071; “ Moduli corrections to gauge and gravitational couplings in four-dimensional superstrings,” Nucl. Phys. **B383** (1992) 109, hep-th/9204030.

[44] I. Antoniadis, E. Gava, K.S. Narain and T.R. Taylor, “Superstring threshhold corrections to Yukawa couplings,” Nucl. Phys **B407** (1993) 706; hep-th/9212045. Note: These versions are different.

[45] I. Antoniadis and T.R. Taylor, “String loop corrections to gauge and Yukawa couplings,” hep-th/9301033.

[46] D. Bailin, A. Love, W. A. Sabra and S. Thomas, “Duality symmetries of threshold corrections in orbifold models,” Phys. Lett. **B320** (1994) 21 ;
“String loop threshold corrections for Z(N) Coxeter orbifolds,” Mod. Phys. Lett. **A9** (1994) 67.

[47] P. Mayr and S. Stieberger, “Threshold corrections to gauge couplings in orbifold compactifications,” Nucl. Phys. **B407** (1993) 725, hep-th/9303017.

[48] V. Kaplunovsky and J. Louis, “On gauge couplings in string theory,” Nucl. Phys. **B444** (1995) 191, hep-th/9502077.

[49] S. Cecotti, P. Fendley, K. Intriligator and C. Vafa, “A new supersymmetric index, ” Nucl. Phys. **B386** (1992) 405, hep-th/9204102; S. Cecotti and C. Vafa, “Ising model and $N = 2$ supersymmetric theories, ” Commun. Math. Phys. **157** (1993) 139, hep-th/9209085.

[50] H. Ooguri and C. Vafa, “Geometry of $N = 2$ strings,” Nucl. Phys. **B361** (1991) 969.

[51] S. Ferrara, C. Kounnas, D. Lüst and F. Zwirner, “Duality-invariant partition functions and automorphic superpotentials for $(2, 2)$ string compactifications,” Nucl. Phys. **B365** (1991) 431.

[52] C. Vafa, “A stringy test of the fate of the conifold,” Nucl. Phys. **B447** (1995) 252, hep-th/9505053.

[53] T. Banks and L. Dixon, “Constraints on string vacua with space-time supersymmetry,” Nucl. Phys. **B307** (1988) 93.

[54] T. Eguchi and A. Taormina, “Character formulas for the $N = 4$ superconformal algebra,” Phys. Lett. **200B** (1988) 315.

[55] S. Ferrara and A. Van Proeyen, “ A theorem on $N = 2$ special Kahler product manifolds,” Class. Quantum Grav. **6** (1989) L243.

[56] V. Kaplunovsky, J. Louis and S. Theisen, “Aspects of duality in $N = 2$ string vacua,” Phys. Lett. **B357** (1995) 71, hep-th/9506110.

[57] A. Klemm, W. Lerche and P. Mayr, “K3-fibrations and Heterotic-Type II string duality, ” Phys. Lett. **B357** (1995) 313, hep-th/9506122.

[58] I. Antoniadis, E. Gava, K. S. Narain and T. R. Taylor, “$N = 2$ Type II- Heterotic duality and higher derivative F-terms, ” hep-th/9507115.

[59] P. Mayr and S. Steiberger, “Moduli dependence of one loop gauge couplings in $(0, 2)$ compactifications,” Phys. Lett. **B355** (1995) 107, hep-th/9504129.

[60] A. Ceresole, R. D’Auria, S. Ferrara and A. Van Proeyen, “On Electromagnetic Duality in Locally Supersymmetric N=2 Yang–Mills Theory,” hep-th/9412200.

[61] E. Calabi and E. Vesentini, Ann. Math. **71** (1960) 472.

[62] R. Gilmore, *Lie Groups, Lie Algebras and some of their Applications,* Wiley-Interscience, New York, 1974.

[63] A. N. Schellekens and N. P. Warner, Phys. Lett. **177B** (1986) 317; Phys. Lett. **181B** (1986) 339; Nucl. Phys. **B287** (1987) 317.

[64] E. Witten, “Elliptic Genera and Quantum Field Theory,” Commun. Math. Phys. **109**(1987)525; “The index of the Dirac operator in loop space,” Proceedings of the conference on elliptic curves and modular forms in algebraic topology, Princeton NJ, 1986.

[65] W. Lerche, “Elliptic index and superstring effective actions,” Nucl. Phys. **B308** (1988) 102.

[66] T. Eguchi, H. Ooguri, A. Taormina and S. K. Yang, “Superconformal algebras and string compactification on manifolds with $SU(n)$ holonomy,” Nucl. Phys. **B315** (1989) 193.

[67] J. P. Derendinger, S. Ferrara, C. Kounnas and F. Zwirner, “On loop corrections to string effective field theories: field-dependent gauge couplings and σ-model anomalies,” Nucl. Phys. **B372** (1992) 145.

[68] G. L. Cardoso and B. A. Ovrut, “A Green-Schwarz mechanism for $D = 4$, $N = 1$ supergravity anomalies,” Nucl. Phys. **B369** (1992) 351; “Coordinate and Kahler sigma model anomalies and their cancellation in string effective field theories,” Nucl. Phys. **B392** (1993) 315, hep-th/9205009.

[69] E. Kiritsis and C. Kounnas, "Infrared Regularization of superstring theory and the one-loop calculation of coupling constants," Nucl. Phys. **B442** (1995) 442, hep-th/9501020.

[70] L. Lewin, Polylogarithms and Associated Functions, North Holland 1981. See eq. 6.7.

[71] R. Borcherds, "Generalized Kac-Moody algebras," Journal of Algebra **115** (1988) 501.

[72] V. G. Kac, *Infinite dimensional Lie algebras,* Cambridge University Press, Cambridge, 1990.

[73] J. Fuchs, *Affine lie algebras and quantum groups,* Cambridge University Press, Cambridge, 1992.

[74] B. Julia, in *Applications of group theory in physics and mathematical physics,* ed. P. Sally et. al. (American Mathematical Society, Providence, 1985).

[75] V. G. Kac, R. V. Moody, and M. Wakimoto, "On E_{10}," In K. Bleuler and M. wener, eds. *Differential geometrical methods in theoretical physics.* Proceedings, NATO advanced research workshop, 16th international conference, Como Kluwer, 1988.

[76] R. W. Gebert and H. Nicolai, "E_{10} for beginners," hep-th/9411188.

[77] I. Frenkel, H. Garland, and G. Zuckerman, "Semi-infinite cohomology and string theory," Proc. Nat. Acad. Sci. **83**(1986) 8442.

[78] see e.g. V. V. Nikulin, "Reflection groups in hyperbolic spaces and the denominator formula for Lorentzian Kac-Moody algebras, alg-geom/9503003.

[79] I. B. Frenkel, J. Lepowsky, and A. Meurman, *Vertex operator algebras and the monster,* Pure and Applied Mathematics Volume 134, Academic Press, San Diego, 1988.

[80] D. Gross, "High energy symmetries of string theory," Phys. Rev. Lett. **60B** (1988) 1229.

[81] B. de Wit and A. Van Proeyen, "Broken sigma-model isometries in very special geometry," Phys. Lett. **293B**(1992)94.

[82] H. Ooguri and C. Vafa, " $N = 2$ heterotic strings, " Nucl. Phys. **B367** (1991) 83.

[83] D. Friedan, E. Martinec, and S. Shenker, "Conformal Invariance, Supersymmetry, and String Theory," Nucl.Phys. **B271** (1986) 93.

[84] H. Yoshii, "On moduli space of $c = 0$ topological conformal field theories," Phys. Lett. **B275** (1992) 70.

[85] S. Nojiri, "$N = 2$ superconformal topological field theory," Phys. Lett. **264B**(1991)57.

[86] N. Berkovits and C. Vafa, "$N = 4$ topological strings, " Nucl. Phys. **B433** (1995) 123, hep-th/9407190.

[87] E. Witten and B. Zwiebach, "Algebraic structures and differential geometry in 2d string theory," Nucl. Phys. **B377**(1992)55;hep-th/9201056

[88] B. H. Lian and G. J. Zuckerman, "New Perspectives on the BRST Algebraic Structure of String Theory," Commun.Math.Phys. **154** (1993) 613, hep-th/9211072.

[89] P. Candelas, X. de la Ossa, A. Font, S. Katz, and D. Morrison, "Mirror Symmetry for Two Parameter Models - I " Nucl. Phys. **B416** (1994) 481, hep-th/9308083.

[90] Work in progress.

[91] G. Aldazabal, A. Font, L. E. Ibáñez and F. Quevedo, "Chains of $N = 2, D = 4$ heterotic/type II duals," hep-th/9510093.

[92] E. Martinec, "Criticality, Catastrophes, and Compactifications," in L. Brink et. al. eds. *Physics and Mathematics of Strings*, V. Knizhnik memorial volume. World Scientific, 1990.

[93] P.S. Aspinwall and D.R. Morrison, "String Theory on K3 Surfaces," hep-th/9404151.

[94] P. Candelas, X. de la Ossa, P. S. Green and L. Parkes, "A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory," Nucl. Phys. **B359** (1991) 21.

[95] M. T. Grisaru, A. van de Ven and D. Zanon, Phys. Lett. **173B** (1986) 423, Nucl. Phys. **B277** (1986) 388, Nucl. Phys. **B277** (1986) 409.

[96] A. Feingold and I. Frenkel, "A Hyperbolic Kac-Moody Algebra and the Theory of Siegel Modular Forms of Genus 2," Math. Ann. **263** (1083) 87.

[97] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, " Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes, " Commun. Math. Phys. **165** (1994) 311, hep-th/9309140.

[98] J. Jorgenson and A. Todorov, "A conjectured analog of Dedekind's eta function for K3 surfaces," Yale preprint.

[99] B. H. Lian and S. T. Yau, "Arithmetic properties of mirror map and quantum coupling, " hep-th/9411234; "Mirror Maps, Modular Relations and Hypergeometric Series I ," hep-th/9506210; "Mirror Maps, Modular Relations and Hypergeometric Series II," hep-th/9507153

[100] P. Horava, "Strings on world sheet orbifolds," Nucl. Phys. **B327** (1989) 461; "Background duality of open string models," Phys. Lett. **B231**(1989)251.

[101] M. Bianchi, G. Pradisi and A. Sagnotti, "Toroidal compactification and symmetry breaking in open string theories," Nucl. Phys. **B376** (1992) 365.

[102] J. Polchinski, "Combinatorics of boundaries in string theory," Phys. Rev. **D50** (1994) 6041, hep-th/9407031; M. B. Green, "A gas of D instantons," Phys. Lett. **B354** (1995) 271, hep-th/9504108.

[103] C. Vafa and E. Witten, "Dual string pairs with $N = 1$ and $N = 2$ supersymmetry in four dimensions, " hep-th/9507050.

[104] E. Witten, "Space-time and topological orbifolds," Phys. Rev. Lett. **61** (1988) 670.

[105] see e.g. W. Lerche, A. N. Schellekens and N. P. Warner, " Lattices and Strings, " Phys. Rep. **177** (1989) 1.

[106] J. Patera, R. T. Sharp and P. Winternitz, J. Math. Phys. **17** (1976) 1972.

A Monster Lie Algebra?

Chapter 30 of “Sphere packing, lattices and groups” by Conway and Sloane, and Adv. in Math. 53 (1984), no. 1, 75–79.

R. E. Borcherds, J. H. Conway, L. Queen and N. J. A. Sloane

We define a remarkable Lie algebra of infinite dimension, and conjecture that it may be related to the Fischer-Griess Monster group.

The idea was mooted in [C-N] that there might be an infinite-dimensional Lie algebra (or superalgebra) L that in some sense “explains” the Fischer-Griess ‘Monster” group M. In this chapter we produce some candidates for L based on properties of the Leech lattice described in [C-S]. These candidates are described in terms of a particular Lie algebra L_∞ of infinite rank.

We first review some of our present knowledge about these matters. It was proved by character calculations in [C-N, p. 317] the centralizer C of an involution of class 2A in the Monster group has a natural sequence of modules affording the head characters (restricted to C). In [K], V. Kac has explicitly constructed these as C-modules. Now that Atkin, Fong and Smith [F], [S] have verified the relevant numerical conjectures of [C-N] for M, we know that these modules can be given the structure of M-modules. More recently, Frenkel, Lepowsky, and Meurman [F-L-M] have given a simple construction for the monster along these lines, but this sheds little light on the conjectures.

Some of the conjectures of [C-N] have analogs in which M is replaced by a compact simple Lie group, and in particular by the Lie group E_8. Most of the resulting statements have now been established by Kac and others. However, it seems that this analogy with Lie groups may not be as close as one would wish, since two of the four conjugacy classes of elements of order 3 in E_8 were shown in [Q] to yield examples of modular functions neither of which are the Hauptmodul for any modular group. This disproves the conjecture made on p. 267 of [K], and is particularly distressing since it was the Hauptmodul property that prompted the discovery of the conjectures in [C-N], and it is this property that gives those conjectures almost all their predictive power.

The properties of the Leech lattice that we shall use stem mostly from the facts about “deep holes” in that lattice reported in [C-S Chapter 23]. Let $w = (0, 1, 2, 3, \ldots, 24 \mid 70)$. The main result of [C-S Chapter 26] is that the subset of vectors r in $II_{25,1}$ for which $r \cdot r = 2$, $r \cdot w = -1$ (the “Leech roots’) is isometric to the Leech lattice, under the metric defined by $d(r, s)^2 = \text{norm}(r - s)$. The main result of [C-S Chapter 27] is that $\text{Aut}(\text{II}_{25,1})$ is obtained by extending the Coxeter subgroup generated by the reflections in these Leech roots by its group of graph automorphisms together with the central inversion -1. It is remarkable that the walls of the fundamental region for this Coxeter group (which correspond one-for-one with the Leech roots) are transitively permuted by the graph automorphisms, which form an infinite group abstractly isomorphic to the group of all automorphisms of the Leech lattice, including translations.

Vinberg [V] shows that for the earlier analogs $II_{9,1}$ and $II_{17,1}$ of $II_{25,1}$ the fundamental regions for the reflection subgroups have respectively 10 and 19 walls, and the graph automorphism groups have orders 1 and 2. For the later analogs $II_{33,1}, \ldots$, there is no “Weyl vector” like w, so it appears that $II_{25,1}$ is very much a unique object.

We can use the vector w to define a root system in $II_{25,1}$. If $v \in II_{25,1}$ then we define the height of v by $-v \cdot w$, and we say that v is positive or negative according as its height is positive or negative. We now define a Kac-Moody Lie algebra L_∞, of infinite dimension and rank, as follows: L_∞ has three generators $e(r)$, $f(r)$, $h(r)$ for each Leech root r, and is presented by the following relations:

$$\begin{aligned}
[e(r), h(s)] &= r \cdot s\ e(r)\ , \\
[f(r), h(s)] &= -r \cdot s\ f(r), \\
[e(r), f(r)] &= h(r), \\
[e(r), f(s)] &= 0, \\
[h(r), h(r)] &= 0 = [h(r), h(s)], \\
e(r)\ \{\mathrm{ad}\ e(s)\}^{1-r\cdot s} &= 0 = f(r)\ \{\mathrm{ad}\ f(s)\}^{1-r\cdot s},
\end{aligned}$$

where r and s are distinct Leech roots. (We have quoted these relations from Moody's excellent survey article [M]. Moody supposes that the number of fundamental roots is finite, but since no argument ever refers to infinitely many fundamental roots at once, this clearly does not matter. See also [K89].)

Then we conjecture that L_∞ provides a natural setting for the Monster, and more specifically that the Monster can be regarded as a subquotient of the automorphism group of some naturally determined subquotient algebra of L_∞.

The main problem is to "cut L_∞ down to size'. Here are some suggestions. A rather trivial remark is that we can replace the Cartan subalgebra H of L_∞ by the homomorphic image obtained by adding the relations

$$c_1\ h(r_1)\ +\ c_2\ h(r_2)\ +\ \ldots\ =\ 0$$

for Leech roots $r_1, r_2, \ldots$, whenever $c_1, c_2, \ldots$ are integers for which

$$c_1 r_1\ +\ c_2 r_2\ +\ \ldots\ =\ 0\ .$$

A more significant idea is to replace L_∞ by some kind of completion allowing us to form infinite linear combinations of the generators, and then restrict to the subalgebra fixed by all the graph automorphisms. The resulting algebra, supposing it can be defined, would almost certainly not have any notion of root system.

Other subalgebras of L_∞ are associated with the holes in the Leech lattice, which are either "deep" holes or "shallow" holes (see [C-S Chapter 23]).

(i) By [C-S Chapter 23], any deep hole corresponds to a Niemeier lattice N, which has a Witt part which is a direct sum of root lattices chosen from the list A_n $(n = 1, 2, \ldots)$, D_n $(n = 4, 5, \ldots)$, E_6, E_7 and E_8. Only 23 particular combinations arise, and we shall take $A_{11} D_7 E_6$ as our standard example. The graph of Leech roots contains a finite subgraph which is the disjoint union of extended Dynkin diagrams corresponding to these Witt components W of N, and so our algebra L_∞ has a subalgebra $L[N]$ which is a direct sum of the Euclidean Lie algebras $E(W)$ corresponding to those components (see [K89]). For example, L_∞ has a subalgebra

$$E(A_{11}) \ + \ E(D_7) \ + \ E(E_6) \ .$$

Each such subalgebra of L_∞ can be extended to a larger subalgebra $L^*(N)$ having one more fundamental root, corresponding to a "glue vector' of the appropriate hole (see [C-S Chapter 25]). In the corresponding graph, the new node is joined to a single special node in each component. The graph for $L^*[A_{11}D_7E_6]$ is shown in Figure 30.1 of [C-S]. (A special node of a connected extended Dynkin diagram is one whose deletion would result in the corresponding ordinary diagram.) These hyperbolic algebras $L^*[N]$, having finite rank, are certainly more manageable than L_∞ itself. Since the 23 Niemeier lattices yield 23 constructions for the Leech lattice ([C-S Chapter 25]), it is natural to ask if we can obtain 23 different constructions for the Monster using the Lie algebras $L[N]$ or $L^*[N]$.

(ii) Each shallow hole in the Leech lattice (see [C-S Chapter 24]) corresponds to a maximal subalgebra of L_∞ of finite rank.

We are making various calculations concerning L_∞ (finding the multiplicities of certain roots via the Weyl-MacDonald-Kac formula, etc.). It is worth noting that these calculations are facilitated by the remarkable recent discovery that the Mathieu group M_{12} is generated by the two permutations

$$t \ \mapsto \ |2t| \ , \qquad t \ \mapsto \ 11 - t \ (\text{mod } 23) \ ,$$

of the set $\{0,1,2,3,4,5,6,7,8,9,10,11\}$, where $|x|$ denotes the unique y in this set for which $y \ \equiv \ \pm x$ (mod 23). (See [C-S Chapter 11]. This discovery arose from the study of properties of various standard playing-card shuffles. We have noticed that there are many other elements of M_{12} which have simple formulae in this "card numbering", for example $t \ \mapsto \ | \ t^3 \ |$.) The simplest transformation (see [C-S Chapter 28]) between the usual Euclidean coordinates for the Leech lattice and its Lorentzian coordinates uses this description of M_{12}.

References.

[C-N] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[C-S] J. H. Conway, N. J. A. Sloane, Sphere packings, lattices and groups. Third edition. Grundlehren der Mathematischen Wissenschaften 290. Springer-Verlag, New York, 1999. ISBN: 0-387-98585-9

[F] P. Fong, Characters arising in the Monster–modular connection. P.S.P.M. 37 (1980), 557–559.

[F-L-M] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[K] V. G. Kac, An elucidation of "Infinite dimensional algebras … and the very strange formula". Adv. in Math. 35 (1980), 264–273.

[K89] V. G. Kac, "Infinite dimensional Lie algebras", third edition, Cambridge University Press, 1990.

[M] R. V. Moody, Root systems of hyperbolic type, Adv. in Math. 33 (1979), 144–160.

[Q] L. Queen, Some relations between finite groups, Lie groups, and modular functions, PhD. dissertation, Univ. of Cambridge, 1980. Modular functions and finite simple groups, PSPM 37 (1980), 561–566. Modular functions arising from some finite groups, MTAC 37 (1981), 547–580.

[S] R. F. Smith, On the head characters of the monster simple group, In "Finite groups–coming of age", edited by J. McKay, Contemp. Math. 45 (1985) p. 303–313.

[V] È. B. Vinberg, Some arithmetical discrete groups in Lobačevskiĭ spaces. Discrete subgroups of Lie groups and applications to moduli (Internat. Colloq., Bombay, 1973), pp. 323–348. Oxford Univ. Press, Bombay, 1975.

Remark added 1998: The Lie algebra of this paper is indeed closely related to the monster simple group. In order to get a well behaved Lie algebra it turns out to be necessary to add some imaginary simple roots to the "Leech roots". This gives the *fake* monster Lie algebra, which contains the Lie algebra of this paper as a large subalgebra. See "The monster Lie algebra", Adv. Math. Vol. 83, No. 1, Sept. 1990, for details (but note that the fake monster Lie algebra is called the monster Lie algebra in this paper). The term "monster Lie algebra" is now used to refer to a certain "$Z/2Z$-twisted" version of the fake monster Lie algebra. The monster Lie algebra is acted on by the monster simple group, and can be used to show that the monster module constructed by Frenkel, Lepowsky, and Meurman satisfies the moonshine conjectures; see "Monstrous moonshine and monstrous Lie superalgebras", Invent. Math. 109, 405-444 (1992).

Preprint typeset in JHEP style - HYPER VERSION

arXiv:1208.4074v2 [hep-th] 3 Apr 2014

Quantum Black Holes, Wall Crossing, and Mock Modular Forms

Atish Dabholkar[1,2], **Sameer Murthy**[3], **and Don Zagier**[4,5]

[1] *Theory Division, CERN, PH-TH Case C01600, CH-1211, 23 Geneva, Switzerland*

[2] *Sorbonne Universités, UPMC Univ Paris 06*
UMR 7589, LPTHE, F-75005, Paris, France

CNRS, UMR 7589, LPTHE, F-75005, Paris, France

[3] *NIKHEF theory group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

[4] *Max-Planck-Institut für Mathematik,Vivatsgasse 7, 53111 Bonn, Germany*

[5] *Collège de France, 3 Rue d'Ulm, 75005 Paris, France*

Emails: `atish@lpthe.jussieu.fr`, `murthy.sameer@gmail.com`, `don.zagier@mpim-bonn.mpg.de`

Abstract: We show that the meromorphic Jacobi form that counts the quarter-BPS states in $\mathcal{N} = 4$ string theories can be canonically decomposed as a sum of a *mock Jacobi form* and an *Appell-Lerch sum.* The quantum degeneracies of single-centered black holes are Fourier coefficients of this mock Jacobi form, while the Appell-Lerch sum captures the degeneracies of multi-centered black holes which decay upon wall-crossing. The *completion* of the mock Jacobi form restores the modular symmetries expected from AdS_3/CFT_2 holography but has a holomorphic anomaly reflecting the non-compactness of the microscopic CFT. For every positive integral value m of the magnetic charge invariant of the black hole, our analysis leads to a special mock Jacobi form of weight two and index m, which we characterize uniquely up to a Jacobi cusp form. This family of special forms and another closely related family of weight-one forms contain almost all the known mock modular forms including the mock theta functions of Ramanujan, the generating function of Hurwitz-Kronecker class numbers, the mock modular forms appearing in the Mathieu and Umbral moonshine, as well as an infinite number of new examples.

Keywords: black holes, mock modular forms, superstrings, wall-crossing.

Contents

1. Introduction **3**
- 1.1 Introduction for mathematicians 3
- 1.2 Introduction for physicists 6
- 1.3 Organization of the paper 12

2. Review of Type-II superstring theory on $K3 \times T^2$ **12**

3. Modular forms in one variable **17**
- 3.1 Basic definitions and properties 18
- 3.2 Quantum black holes and modular forms 21

4. Jacobi forms **23**
- 4.1 Basic definitions 23
- 4.2 Theta expansion and Taylor expansion 24
- 4.3 Example: Jacobi forms of index 1 27
- 4.4 Hecke-like operators 29
- 4.5 Quantum black holes and Jacobi forms 32

5. Siegel modular forms **33**
- 5.1 Definitions and examples of Siegel modular forms 33
- 5.2 Quantum black holes and Siegel modular forms 36

6. Walls and contours **37**

7. Mock modular forms **40**
- 7.1 Pure mock modular forms 40
- 7.2 Mock Jacobi forms 44
- 7.3 Mock modular forms: the general case 48
- 7.4 Superconformal characters and mock modular forms 50

8. From meromorphic Jacobi forms to mock modular forms **51**
- 8.1 The Fourier coefficients of a meromorphic Jacobi form 52
- 8.2 The polar part of φ (case of simple poles) 53
- 8.3 Mock modularity of the Fourier coefficients 57
- 8.4 The case of double poles 61
- 8.5 Examples 64

9. Special mock Jacobi forms **67**
- 9.1 The simplest meromorphic Jacobi forms 67
- 9.2 Choosing the function φ for weight 2 and small index 69
- 9.3 Further data and formulation of main results for the weight 2 family 75
- 9.4 Optimal choice of the function $\Phi_{2,m}$ 81
- 9.5 Observations on the weight one family, integrality, and positivity 86
- 9.6 Higher weights 95

10. Structure theorems for Jacobi and mock Jacobi forms **98**
- 10.1 Relations among the polar coefficients of weak Jacobi forms 98
- 10.2 Choosing optimal versions of weak elliptic forms 102
- 10.3 The residues of mock Jacobi forms 107
- 10.4 Remaining proofs of theorems from §9 109

11. Quantum black holes and mock modular forms **115**
- 11.1 Mock Jacobi forms, immortal black holes, and AdS_2 holography 115
- 11.2 Meromorphy, mock modularity, and noncompactness 120
- 11.3 The M-Theory limit and AdS_3 holography 122
- 11.4 Open problems and comments 126

A. Appendix: Tables of special mock Jacobi forms **130**
- A.1 Table of $\mathcal{Q}_M$ (weight 2 case) 130
- A.2 Table of $\Phi^0_{1,m}$ of optimal growth 134
- A.3 Table of $\mathcal{Q}_M$ (weight 1 case) 139

B. Bibliography **144**

"My dream is that I will live to see the day when our young physicists, struggling to bring the predictions of superstring theory into correspondence with the facts of nature, will be led to enlarge their analytic machinery to include not only theta-functions but mock theta-functions . . . But before this can happen, the purely mathematical exploration of the mock-modular forms and their mock-symmetries must be carried a great deal further."

Freeman Dyson (1987 Ramanujan Centenary Conference)

1. Introduction

The quantum degeneracies associated with a black hole horizon are of central importance in quantum gravity. The counting function for these degeneracies for a large class of black holes in string theory is expected to be modular from the perspective of holography. However, in situations with wall-crossing, there is an apparent loss of modularity and it is far from clear if and how such a counting function can be modular. In the context of quarter-BPS black holes in $\mathcal{N} = 4$ supersymmetric theories, we develop the required analytic machinery that provides a complete answer to this question which leads naturally to the mathematics of mock modular forms. We present a number of new mathematical results motivated by but independent of these physical considerations.

Since this paper is of possible interest to both theoretical physicists (especially string theorists) and theoretical mathematicians (especially number theorists), we give two introductions in their respective dialects.

1.1 Introduction for mathematicians

In the quantum theory of black holes in the context of string theory, the physical problem of counting the dimensions of certain eigenspaces ("the number of quarter-BPS dyonic states of a given charge") has led to the study of Fourier coefficients of certain *meromorphic* Siegel modular forms and to the question of the modular nature of the corresponding generating functions. Using and refining results of S. Zwegers [130], we show that these generating functions belong to the recently discovered class of functions called *mock modular forms*.

Since this notion is still not widely known, it will be reviewed in some detail (in §7.1). Very roughly, a mock modular form of weight k (more precisely, "pure" mock modular forms; we will also introduce a somewhat more general notion of "mixed" mock modular forms in §7.3) is a holomorphic function f in the upper half plane to which is associated a holomorphic modular form g of weight $2-k$, called the "shadow" of f, such that the sum of f and a suitable non-holomorphic integral of g transforms like a holomorphic modular form of weight k. Functions of this type occur in several contexts in mathematics and mathematical physics: as certain q-hypergeometric series (like Ramanujan's original mock theta functions), as generating functions of class numbers of imaginary quadratic fields [126], or as characters of extended superconformal algebras [99], with special cases of the last class being conjecturally related to the Mathieu group M_{24} [52]. They also arise, as was shown by Zwegers in his thesis [130], as the Fourier coefficients of meromorphic Jacobi forms. It is this last occurrence which is at the origin of the connection to black hole physics, because the Fourier coefficients of meromorphic Jacobi forms have the same wall-crossing behavior as that exhibited by the degeneracies of BPS states.

The specific meromorphic Jacobi forms that are of interest for the black hole counting problem are the Fourier-Jacobi coefficients $\psi_m(\tau, z)$ of the meromorphic Siegel modular form

$$\frac{1}{\Phi_{10}(\Omega)} \;=\; \sum_{m=-1}^{\infty} \psi_m(\tau, z)\, p^m\,, \qquad \left(\Omega = \begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix},\ p = e^{2\pi i \sigma}\right), \tag{1.1}$$

the reciprocal of the Igusa cusp form of weight 10, which arises as the partition function of quarter-BPS dyons in the type II compactification on the product of a $K3$ surface and an elliptic curve [51, 109, 42]. These coefficients, after multiplication by the discriminant function $\Delta(\tau)$, are meromorphic Jacobi forms of weight 2 with a double pole at $z = 0$ and no others (up to translation by the period lattice).

The new mathematical results of the paper are contained in Sections 8, 9 and 10. In §8, extending the results of [130], we show that any meromorphic Jacobi form $\varphi(\tau, z)$ having poles only at torsion points $z = \alpha\tau + \beta$ ($\alpha,\, \beta \in \mathbb{Q}$) has a canonical decomposition into two pieces $\varphi^{\mathrm{F}}(\tau, z)$ and $\varphi^{\mathrm{P}}(\tau, z)$, called its "finite" part and "polar" part, respectively, of which the first is a finite linear combination of classical theta series with mock modular forms as coefficients, and the second is an elementary expression that is determined completely by the poles of φ. Again using the results of Zwegers, we give explicit formulas for the polar part φ^P for all φ having only simple or double poles. In the particular case of ψ_m, the polar part is given by the formula

$$\psi_m^{\mathrm{P}}(\tau, z) \;=\; \frac{p_{24}(m+1)}{\Delta(\tau)}\, \mathcal{A}_{2,m}(\tau, z) \qquad \left(q = e^{2\pi i \tau}\,,\ y = e^{2\pi i z}\right), \tag{1.2}$$

where $\Delta(\tau)$ is the Ramanujan discriminant function, $p_{24}(m+1)$ the coefficient of q^m in $\Delta(\tau)^{-1}$, and $\mathcal{A}_{2,m}(\tau, z)$ the elementary function (Appell-Lerch sum)

$$\mathcal{A}_{2,m}(\tau, z) \;=\; \sum_{s \in \mathbb{Z}} \frac{q^{ms^2+s} y^{2ms+1}}{(1 - q^s y)^2}\,. \tag{1.3}$$

Note that $\mathcal{A}_{2,m}$ exhibits wall-crossing: for $0 < \mathrm{Im}(z) < \mathrm{Im}(\tau)$ it has the Fourier expansion

$$\mathcal{A}_{2,m}(\tau, z) \;=\; \sum_{\substack{r \geq \ell > 0 \\ r \equiv \ell \pmod{2m}}} \ell\, q^{\frac{r^2 - \ell^2}{4m}}\, y^r\,, \tag{1.4}$$

but the Fourier expansions are different in other strips $n < \mathrm{Im}(z)/\mathrm{Im}(\tau) < n + 1$. This wall-crossing is at the heart of both the mathematical and the physical theories. On the mathematical side it explains the "mockness" of the finite part ψ_m^{F} of ψ_m. On the physical side it has an interpretation in terms of counting two-centered black holes, with the integer ℓ in (1.4) being the dimension of the $SU(2)$ multiplet with angular momentum $(\ell - 1)/2$ contained in the electromagnetic field produced by the two centers.

Starting in §9 we focus attention on two particular classes of functions $\{\varphi_{1,m}\}$ and $\{\varphi_{2,m}\}$, where $\varphi_{k,m}$ is a meromorphic Jacobi form of weight $k \in \{1, 2\}$ and index $m \in \mathbb{N}$ having singular part $(2\pi iz)^{-k} + \mathrm{O}(1)$ as $z \to 0$ and no other poles except the translates of this one by the period lattice. These functions in the case $k = 2$ are related to the Fourier coefficients ψ_m defined in (1.1) by

$$\psi_m(\tau, z) \;=\; \frac{p_{24}(m+1)}{\Delta(\tau)}\,\varphi_{2,m}(\tau, z) \,+\, (\text{weakly holomorphic Jacobi form}). \tag{1.5}$$

The polar part of $\varphi_{2,m}$ is the function $\mathcal{A}_{2,m}$ defined above, and that of $\varphi_{1,m}$ an Appell-Lerch sum $\mathcal{A}_{1,m}$ with a similar definition (eq. (9.6)), but the functions themselves are not unique, since we can add to them any (weak) Jacobi forms of weight k and index m without changing the defining property. This will change the mock Jacobi form $\Phi_{k,m} = \varphi^{\mathrm{F}}_{k,m}$ in the corresponding way. Much of our analysis concerns finding "optimal" choices, meaning choices for which the functions $\Phi_{k,m}$ have poles of as small an order as possible at infinity and consequently Fourier coefficients whose asymptotic growth is as small as possible. Our main results concern the case $k = 2$, and are as follows:

1. If m is a prime power, then the functions $\Phi_{2,m}$ can be chosen to be strongly holomorphic (i.e., with Fourier expansions containing only monomials $q^n y^r$ with $4nm - r^2 \geq 0$). Their Fourier coefficients then have polynomial growth and are given by explicit linear combinations of class numbers of imaginary quadratic fields.
2. More generally, for arbitrary m, the function $\Phi_{2,m}$ can be chosen to be the sum of the images under appropriate "Hecke-like operators" of special mock Jacobi forms $\mathcal{Q}_M$, where M ranges over all divisors of m having an even number of distinct prime factors. These summands are the eigencomponents of $\Phi_{2,m}$ with respect to the Atkin-Lehner-like operators W_{m_1} $(m_1|m,\ (m_1, m/m_1) = 1)$ acting on the space of mock Jacobi forms of index m.
3. The mock Jacobi form $\mathcal{Q}_1$ is the generating function of class numbers of imaginary quadratic fields, and is strongly holomorphic. The other special mock Jacobi forms $\mathcal{Q}_M$ $(M = 6, 10, 14, \ldots)$ can be chosen to have "optimal growth" (meaning that their Fourier expansions contain only monomials $q^n y^r$ with $4nm - r^2 \geq -1$). Their Fourier coefficients then grow like $e^{\pi\sqrt{\Delta}/m}$ as $\Delta := 4nm - r^2$ tends to infinity.
4. One can also choose $\Phi_{2,m}(\tau, z)$ for arbitrary $m \geq 1$ to be of the form $\sum \Phi^0_{2,m/d^2}(\tau, dz)$, where d ranges over positive integers with $d^2|m$ and each $\Phi^0_{2,m/d^2}$ has optimal growth.
5. There are explicit formulas for the polar coefficients (those with $\Delta < 0$) of $\Phi^0_{2,m}$ and $\mathcal{Q}_M$. In particular, the coefficient of $q^n y^r$ in $\mathcal{Q}_M$ vanishes if $4nM - r^2 = 0$ and equals $\pm\varphi(M)/12$ if $4nM - r^2 = -1$, where $\varphi(M) = \prod_{p|M}(p-1)$ is the Euler φ-function of M.

The proofs of these properties are contained in §10, which gives a detailed description (based on difficult results on Jacobi forms proved in [114, 115]) of the way that the space of holomorphic

Jacobi forms is contained in the spaces of weak or of weakly holomorphic Jacobi forms. This analysis contains several surprises, such as the result (Theorem 9.4) that for all $m \geq 1$ the space of holomorphic Jacobi forms of index m has codimension exactly 1 in the space of Jacobi forms of index m with optimal growth.

For $\varphi_{1,m}$ the description is much less complete. There is still a result like 2., with M now ranging over the divisors of m with an *odd* number of distinct prime factors, but it is no longer possible in general to choose the $\mathcal{Q}_M$ to have optimal growth. The few cases where this is possible turn out to be related to mock theta functions that have played a prominent role in the past. Thus $\mathcal{Q}_{30}$ and $\mathcal{Q}_{42}$ are essentially equal to the mock theta functions of "order 5" and "order 7" in Ramanujan's original letter to Hardy, and $\mathcal{Q}_2$ and several of the other $\mathcal{Q}_M$ are related to the "Mathieu moonshine" and "Umbral moonshine" story [52, 25]. The mock Jacobi forms $\mathcal{Q}_M$ for the $k = 2$ case are also related to special mock theta functions, now of weight 3/2, e.g. $\mathcal{Q}_6$ is the mock theta function with shadow $\eta(\tau)$ given in [129].

The applications of the results (for the $k = 2$ case, via the fact that ψ_m is the sum of $p_{24}(m+1)\varphi_{2,m}/\Delta$ and a weakly holomorphic Jacobi form of weight -10 and index m) to the original physics problem are explained in §11 and in the following "second introduction."

1.2 Introduction for physicists

The microscopic quantum description of supersymmetric black holes in string theory usually starts with a brane configuration of given charges and mass at weak coupling, which is localized at a single point in the noncompact spacetime. One then computes an appropriate indexed partition function in the world-volume theory of the branes, which from the perspective of enumerative geometry computes topological invariants such as the Donaldson-Thomas invariants. At strong coupling, the brane configuration gravitates and the indexed partition function is expected to count the microstates of the corresponding macroscopic gravitational configurations. Assuming that the gravitational configuration is a single-centered black hole then gives a way to obtain a statistical understanding of the entropy of the black hole in terms of its microstates, in accordance with the Boltzmann relation[1].

One problem that one often encounters is that the macroscopic configurations are no longer localized at a point and include not only a single-centered black hole of interest but also several multi-centered ones [46, 47, 11, 48]. Moreover, the indexed degeneracy of the multi-centered configurations typically jumps upon crossing walls of marginal stability in the moduli space where the multi-centered configuration breaks up into its single-centered constituents. These jumps are referred to as the 'wall-crossing phenomenon'.

If one is interested in the physics of the horizon or the microstates of a single black hole, the multi-centered configurations and the associated wall-crossings are thus something of a nuisance.

---

[1] Under certain conditions the index equals the absolute number [106, 107, 37].

It is desirable to have a mathematical characterization that isolates the single-centered black holes directly at the microscopic level. One distinguishing feature of single-centered black holes is that they are '*immortal*' in that they exist as stable quantum states for all values of the moduli and hence their degeneracy does not exhibit the wall-crossing phenomenon. We will use this property later to define the counting function for the immortal black holes.

The wall-crossing phenomenon raises important conceptual questions regarding the proper holographic formulation in this context. In many cases, the black hole can be viewed as an excitation of a black string. The near horizon geometry of a black string is AdS_3 which is expected to be holographically dual to a two-dimensional conformal field theory CFT_2. The conformal boundary of Euclidean AdS_3 is a 2-torus with a complex structure parameter τ, and the physical partition function of AdS_3 and of the boundary CFT_2 is a function of τ. The $SL(2,\mathbb{Z})$ transformations of τ can be identified geometrically with global diffeomorphisms of the boundary of AdS_3 space. The partition function is expected to have good modular properties under this geometric symmetry. This symmetry is crucial for the Rademacher-type expansions of the black hole degeneracies for understanding the quantum entropy of these black holes via holography [106, 107, 35, 49, 44, 89, 117, 95, 90, 36]. Implementing the modular symmetries presents several subtleties in situations when there is wall-crossing.

The wall-crossing phenomenon has another important physical implication for the invariance of the spectrum under large gauge transformations of the antisymmetric tensor field. Large gauge transformations lead to the 'spectral flow symmetry' of the partition function of the black string [44]. Since these transformations act both on the charges and the moduli, degeneracies of states with a charge vector Γ at some point ϕ in the moduli space get mapped to the degeneracies of states with charge vector Γ' at some other point ϕ' in the moduli space. Typically, there are many walls separating the point ϕ' and the original point ϕ. As a result, the degeneracies extracted from the black string at a *given* point ϕ in the moduli space do not exhibit the spectral-flow symmetry. On the other hand, the spectrum of immortal black holes is independent of asymptotic moduli and hence must exhibit the spectral-flow symmetry. This raises the question as to how to make the spectral-flow symmetry manifest for the degeneracies of immortal black holes in the generic situation when there is wall-crossing.

With these motivations, our objective will be to isolate the partition functions of the black string associated with immortal black holes and investigate their transformation properties under the boundary modular group and large gauge transformations. More precisely, we would like to investigate the following four questions.

1. Can one define a *microscopic* counting function that cleanly isolates the microstates of immortal black holes from those of the multi-centered black configurations?
2. What are the modular properties of the counting function of immortal black holes when the asymptotic spectrum exhibits the wall-crossing phenomenon?

3. Can this counting function be related to a quantity that is properly modular as might be expected from the perspective of near-horizon AdS_3/CFT_2 holography?
4. Can one define a partition function of the immortal black holes that manifestly exhibits the spectral-flow symmetry resulting from large gauge transformations?

The main difficulties in answering these questions stem from the complicated moduli dependence of the black hole spectrum which is often extremely hard to compute. To address the central conceptual issues in a tractable context, we consider the compactification of Type-II on $K3 \times T^2$ with $\mathcal{N} = 4$ supersymmetry in four dimensions. The spectrum of quarter-BPS dyonic states in this model is exactly computable [51, 60, 109, 110, 85, 42] and by now is well understood at *all* points in the moduli space [105, 31, 26] and for *all* possible duality orbits [31, 7, 8, 6, 33]. Moreover, as we will see, this particular model exhibits almost all of the essential issues that we wish to address. The $\mathcal{N} = 4$ black holes have the remarkable property that even though their spectrum is moduli-dependent, the partition function itself is moduli-independent. The entire moduli dependence of the black hole degeneracy is captured by the moduli dependence of the choice of the Fourier contour [105, 31, 26]. Moreover, the only multi-centered configurations that contribute to the supersymmetric index of quarter-BPS states are the ones with only two centers, each of which is half-BPS [38]. Consequently, the only way the index can jump at a wall is by the decay of a two-centered configuration into its half-BPS constituents [105, 31, 26]; this is a considerable simplification compared to the general $\mathcal{N} = 2$ case where more complicated multi-centered decays are possible. These features make the $\mathcal{N} = 4$ case much more tractable.

The number of microstates of quarter-BPS dyonic states for the above-mentioned compactification is given by a Fourier coefficient of a meromorphic Jacobi form $\psi_m(\tau, z)$ with a moduli-dependent contour. The partition function (1.1) referred to earlier is the generating function for these meromorphic Jacobi forms. Using this simplicity of the moduli dependence and the knowledge of the exact spectrum, it is possible to give very precise answers to the above questions in the $\mathcal{N} = 4$ framework, which turn out to naturally involve mock modular forms.

1. One can define a holomorphic function for counting the microstates of immortal black holes[2] as a Fourier coefficient of the partition function of the black string for a specific choice of the Fourier contour [105, 31, 26]. The contour corresponds to choosing the asymptotic moduli of the theory in the attractor region of the single-centered black hole.
2. Because the asymptotic counting function is a *meromorphic* Jacobi form, the near horizon counting function of immortal black holes is a *mock* modular form in that it fails to be

---

[2] We will use the terms 'immortal' and 'single-centered' interchangeably. In general, the moduli-independent 'immortal' degeneracies can receive contributions not only from single black holes but also from scaling solutions [48]. They are not expected to contribute to the $\mathcal{N} = 4$ index that we consider [38]. In addition, there can be 'hair' degrees of freedom [10, 75], which are degrees of freedom localized outside the black hole horizon that carry part of the charge of the black hole. In frames where the black hole is represented entirely in terms of D-branes, such hair modes are expected to be absent.

modular but in a very specific way. The failure of modularity is governed by a *shadow*, which is given in terms of another holomorphic modular form.

3. Given a mock modular form and its shadow, one can define its *completion* which is a nonholomorphic modular form. The failure of holomorphy can be viewed as a 'holomorphic anomaly' which is also governed by the shadow.
4. The partition function of immortal black holes with manifest spectral-flow invariance is a mock Jacobi form – a new mathematical object defined and elaborated upon in §7.2.

The main physical payoff of the mathematics of mock modular forms in this context is the guarantee that one can still define a completion as in (3) which *is* modular albeit nonholomorphic. As mentioned earlier, the modular transformations on the τ parameter can be identified with global diffeomorphisms of the boundary of the near horizon AdS_3. This connection makes the mathematics of mock modular forms physically very relevant for AdS_3/CFT_2 holography in the presence of wall-crossing and holomorphic anomalies. The required mathematical results concerning mock modular forms are developed in sections §7, §8, §9, and §10.

To orient the physics reader, we summarize the essential conclusions of this mathematical investigation from the perspective of the questions posed above.

1. **Decomposition**: Given an asymptotic counting function, the degeneracies of single-centered black holes can be isolated using the results in §8. Applying Theorem 8.3 to the meromorphic Jacobi form $\psi_m(\tau, z)$ gives a unique decomposition

   $$\psi_m(\tau, z) = \psi_m^{\mathrm{F}}(\tau, z) \, + \, \psi_m^{\mathrm{P}}(\tau, z) \, , \tag{1.6}$$

   such that $\psi_m^{\mathrm{P}}(\tau, z)$ is a simple function (1.2) with the same pole structure in z as $\psi_m(\tau, z)$ and $\psi_m^{\mathrm{F}}(\tau, z)$ has no poles. The elegant decomposition (1.6) is motivated partly by the choice of 'attractor contour' for single-centered black holes and and has a direct physical interpretation: $\psi_m(\tau, z)$ is the counting function of all asymptotic states including both single and multi-centered configurations, $\psi_m^{\mathrm{F}}(\tau, z)$ is the counting function of immortal black holes, whereas $\psi_m^{\mathrm{P}}(\tau, z)$ is the counting function of multi-centered black holes.

   Since both $\psi_m(\tau, z)$ and $\psi_m^{\mathrm{P}}(\tau, z)$ have poles in z, their Fourier coefficients depend on the choice of the contour which in turn depends on the moduli. On the other hand, the Fourier coefficients of $\psi_m^{\mathrm{F}}(\tau, z)$ are unambiguously defined without any contour or moduli dependence. This is what is expected for immortal black holes.

2. **Modular Completion**: The immortal counting function $\psi_m^{\mathrm{F}}(\tau, z)$ defined by the decomposition (1.6) is not modular. However, theorem 8.3 ensures that by adding a specific nonholomorphic function to $\psi_m^{\mathrm{F}}(\tau, z)$, one can obtain its nonholomorphic completion

$\widehat{\psi}^{\mathrm{F}}_m(\tau,z)$ which *is* modular and transforms as a Jabobi form. The failure of holomorphy of the completion $\widehat{\psi}^{\mathrm{F}}_m(\tau,z)$ is given by the equation

$$\tau_2^{3/2}\,\frac{\partial}{\partial\overline{\tau}}\,\widehat{\psi}^{\mathrm{F}}_m(\tau,z) \;=\; \sqrt{\frac{m}{8\pi i}}\,\frac{p_{24}(m+1)}{\Delta(\tau)}\sum_{\ell\,\mathrm{mod}\,(2m)}\overline{\vartheta_{m,\ell}(\tau,0)}\,\vartheta_{m,\ell}(\tau,z)\,. \tag{1.7}$$

Hence the counting function of immortal black holes has a hidden modular symmetry and more specifically is a mock Jacobi form as defined in §7.2. This is one of our main physics results and is described in more detail in §11.

3. **Holomorphic anomaly**: The completion is a natural object to be identified with the indexed partition function of the superconformal field theory $SCFT_2$ dual to a single-centered AdS_3, which is expected to be modular.

   From this perspective, the equation (1.7) can be viewed as a holomorphic anomaly and can in fact be taken as a defining property of a mock Jacobi form for physics applications. Naively, an indexed partition function is expected to be holomorphic because of a cancellation between right-moving bosons and fermions as for the elliptic genus [123]. However, if the $SCFT_2$ is noncompact, then the spectrum is continuous and this naive reasoning may fail leading to an ‘anomaly’. The holomorphic anomaly can then arise as a consequence of the fact that for the right-movers in a noncompact $SCFT$, the density of states of bosons and fermions may be slightly different [119, 4] and may not precisely cancel. The detailed connection between the holomorphic anomaly and the noncompactness of the $SCFT_2$ in this context needs to be understood better from a path-integral perspective.

4. **Optimality**: The Fourier coefficients of $\psi^{\mathrm{F}}_m(\tau,z)$ grow exponentially rapidly as expected for a counting function of black hole degeneracies. It is clear from the anomaly equation (1.7) that if we add a holomorphic true Jacobi form to $\psi^{\mathrm{F}}_m(\tau,z)$ with the same weight and index, it will still admit a modular completion satisfying the same anomaly equation. This raises the question whether for a given holomorphic anomaly there is an ‘optimal’ mock Jacobi form whose Fourier coefficients grow as slowly as possible. The answer to this question (for the functions $\varphi_{2,m}$ related to ψ_m by (1.5)) is in the affirmative but is subtle and requires several new results in the theory of Jacobi forms developed in §10, motivated by and explaining the numerical experiments and observations described in §9.

   A practical implication of such an optimal choice is that the leading contribution to the black hole entropy is then determined essentially by a Fourier coefficient of a true Jacobi form. One can thus apply the familiar Cardy formula and the Rademacher expansion of the Fourier coefficients of true modular forms for the leading answer. There will be exponentially subleading corrections to this leading answer coming from the optimal mock Jacobi form. A Rademacher expansion for these corrections requires a generalization [13, 16] applicable for mock rather than true modular forms.

5. **Examples**: Modular forms with slowly growing Fourier coefficients are mathematically particularly interesting and many of the best-known examples of mock modular forms share this property. An 'optimal' choice thus enables us in many cases to obtain a very explicit expression for the immortal counting function $\psi_m^{\mathrm{F}}(\tau, z)$ in terms of these known mock modular forms. For example, for all m prime, the optimal mock part of $\psi_m^{\mathrm{F}}(\tau, z)$ can be expressed in terms of the generating function of Hurwitz-Kronecker class numbers (see §9.2). On the other hand, for a nonprime m our analysis leads to new mock modular forms with very slowly growing Fourier coefficients.

   The functions $\psi_m(\tau, z)$ that arise in the black hole problem have a double pole at $z = 0$ and its translates, but our methods can also be applied to a related class of functions with just a single pole at $z = 0$. This leads to a second infinite family of mock modular forms, this time of weight 1/2. (In the double pole case, the mock Jacobi forms had weight 2 and their coefficients were mock modular forms of weight 3/2. In both cases, the coefficients are in fact mock theta functions, i.e., mock modular forms whose shadows are unary theta series.) Unlike the first case, where we found that the mock modular forms occurring can always be chosen to have at most simple poles at the cusps, the pole order here is much greater in general, and there are only a handful of examples having only simple poles. It is remarkable that these include essentially all the most prominent examples of mock theta functions, including the original ones of Ramanujan, the mock theta function conjecturally related to the Mathieu group M_{24} [52] and the functions arising in the umbral moonshine conjecture [24].

Modular symmetries are very powerful in physics applications because they relate strong coupling to weak coupling, or high temperature to low temperature. The hidden modular symmetry of mock modular forms is therefore expected to be useful in diverse physics contexts. As mentioned above, mock modularity of the counting function in the present context of black holes is a consequence of meromorphy of the asymptotic counting function which in turn is a consequence of noncompactness of the target space of the microscopic SCFT. Now, conformal field theories with a noncompact target space occur naturally in several physics contexts. For example, a general class of four-dimensional BPS black holes is obtained as a supersymmetric D-brane configuration in Type-II compactification on a Calabi-Yau three-fold X_6. In the M-theory limit, these black holes can be viewed as excitations of the MSW black string [87, 92]. The microscopic theory describing the low energy excitations of the MSW string is the $(0, 4)$ MSW SCFT. The target space of this SCFT does not necessarily have to be compact in which case the considerations of this paper will apply. Very similar objects [126, 73] have already made their appearance in the context of topological supersymmetric Yang-Mills theory on $\mathbb{CP}^2$ [120]. Other examples include the theory of multiple M5-branes [2] quantum Liouville theory and E-strings [91], and the $SL(2,\mathbb{R})/U(1)$ SCFT [54, 119, 4] where the CFT is noncompact. The appearance of a holomorphic anomaly in the regularized Poincaré series for the elliptic genus of CFT_2 was noted in [89] in the context of the $\mathrm{AdS}_3/\mathrm{CFT}_2$ correspondence.

We expect that the general framework of mock modular forms and Jacobi forms developed in this paper is likely to have varied physical applications in the context of non-compact conformal field theories, wall-crossings in enumerative geometry, and recently formulated Mathieu and umbral moonshine conjectures [52, 24].

1.3 Organization of the paper

In §2, we review the physics background concerning the string compactification on $K3 \times T^2$ and the classification of BPS states corresponding to the supersymmetric black holes in this theory. In sections §3, §4, and §5, we review the basic mathematical definitions of various types of classical modular forms (elliptic, Jacobi, Siegel) and illustrate an application to the physics of quantum black holes in each case by means of an example. In §6, we review the moduli dependence of the Fourier contour prescription for extracting the degeneracies of quarter-BPS black holes in the $\mathcal{N} = 4$ theory from the partition function which is a meromorphic Siegel modular form. In §8, we refine results due to Zwegers to show that any meromorphic Jacobi form with poles only at the sub-multiples of the period lattice can be decomposed canonically into two pieces, one of which is a mock Jacobi form. We give explicit formulas for this decomposition in the case when the poles have at most order 2, and again give several examples. In §9 we give a detailed description of the experimental results for the Fourier coefficients of the two families of mock Jacobi forms $\{\Phi_{2,m}\}$ and $\{\Phi_{1,m}\}$, and formulate the main mathematical results concerning them. The proofs of these results are given in §10, after we have formulated and proved a number of structural results about holomorphic and weakly holomorphic Jacobi forms that are new and may be of independent interest. In §11, we apply these results in the physical context to determine the mock Jacobi form that counts the degeneracies of single-centered black holes and discuss the implications for AdS_2/CFT_1 and AdS_3/CFT_2 holography.

2. Review of Type-II superstring theory on $K3 \times T^2$

Superstring theories are naturally formulated in ten-dimensional Lorentzian spacetime $\mathcal{M}_{10}$. A 'compactification' to four-dimensions is obtained by taking $\mathcal{M}_{10}$ to be a product manifold $\mathbb{R}^{1,3} \times X_6$ where X_6 is a compact Calabi-Yau threefold and $\mathbb{R}^{1,3}$ is the noncompact Minkowski spacetime. We will focus in this paper on a compactification of Type-II superstring theory when X_6 is itself the product $X_6 = K3 \times T^2$. A highly nontrivial and surprising result from the 90s is the statement that this compactification is quantum equivalent or 'dual' to a compactification of heterotic string theory on $T^4 \times T^2$ where T^4 is a four-dimensional torus [74, 124]. One can thus describe the theory either in the Type-II frame or the heterotic frame.

The four-dimensional theory in $\mathbb{R}^{1,3}$ resulting from this compactification has $\mathcal{N} = 4$ supersymmetry[3]. The massless fields in the theory consist of 22 vector multiplets in addition to the supergravity multiplet. The massless moduli fields consist of the S-modulus λ taking values in the coset

$$SL(2,\mathbb{Z})\backslash SL(2,\mathbb{R})/O(2,\mathbb{R}), \tag{2.1}$$

and the T-moduli μ taking values in the coset

$$O(22,6,\mathbb{Z})\backslash O(22,6,\mathbb{R})/O(22,\mathbb{R}) \times O(6,\mathbb{R}). \tag{2.2}$$

The group of discrete identifications $SL(2,\mathbb{Z})$ is called the S-duality group. In the heterotic frame, it is the electro-magnetic duality group [101, 102], whereas in the type-II frame, it is simply the group of area-preserving global diffeomorphisms of the T^2 factor. The group of discrete identifications $O(22,6,\mathbb{Z})$ is called the T-duality group. Part of the T-duality group $O(19,3,\mathbb{Z})$ can be recognized as the group of geometric identifications on the moduli space of K3; the other elements are stringy in origin and have to do with mirror symmetry.

At each point in the moduli space of the internal manifold $K3 \times T^2$, one has a distinct four-dimensional theory. One would like to know the spectrum of particle states in this theory. Particle states are unitary irreducible representations, or supermultiplets, of the $\mathcal{N} = 4$ superalgebra. The supermultiplets are of three types which have different dimensions in the rest frame. A long multiplet is 256-dimensional, an intermediate multiplet is 64-dimensional, and a short multiplet is 16-dimensional. A short multiplet preserves half of the eight supersymmetries (*i.e.* it is annihilated by four supercharges) and is called a half-BPS state; an intermediate multiplet preserves one quarter of the supersymmetry (*i.e.* it is annihilated by two supercharges), and is called a quarter-BPS state; and a long multiplet does not preserve any supersymmetry and is called a non-BPS state. One consequence of the BPS property is that the spectrum of these states is 'topological' in that it does not change as the moduli are varied, except for jumps at certain walls in the moduli space [125].

An important property of a BPS states that follows from the superalgebra is that its mass is determined by its charges and the moduli [125]. Thus, to specify a BPS state at a given point in the moduli space, it suffices to specify its charges. The charge vector in this theory transforms in the vector representation of the T-duality group $O(22,6,\mathbb{Z})$ and in the fundamental representation of the S-duality group $SL(2,\mathbb{Z})$. It is thus given by a vector $\Gamma^{\alpha I}$ with integer entries

---

[3] This supersymmetry is a super Lie algebra containing $ISO(1,3) \times SU(4)$ as the even subalgebra where $ISO(1,3)$ is the Poincaré symmetry of the $\mathbb{R}^{1,3}$ spacetime and $SU(4)$ is an internal symmetry usually referred to as R-symmetry. The odd generators of the superalgebra are called supercharges. With $\mathcal{N} = 4$ supersymmetry, there are eight complex supercharges which transform as a spinor of $ISO(1,3)$ and a fundamental of $SU(4)$.

$$\Gamma^{\alpha I} \;=\; \begin{pmatrix} N^I \\ M^I \end{pmatrix} \qquad \text{where} \qquad \alpha = 1,2\,; \quad I = 1,2,\ldots 28 \tag{2.3}$$

transforming in the $(2,28)$ representation of $SL(2,\mathbb{Z}) \times O(22,6,\mathbb{Z})$. The vectors N and M can be regarded as the quantized electric and magnetic charge vectors of the state respectively. They both belong to an even, integral, self-dual lattice $\Pi^{22,6}$. We will assume in what follows that $\Gamma = (N,M)$ in (2.3) is primitive in that it cannot be written as an integer multiple of (N_0, M_0) for N_0 and M_0 belonging to $\Pi^{22,6}$. A state is called purely electric if only N is non-zero, purely magnetic if only M is non- zero, and dyonic if both M and N are non-zero.

To define S-duality transformations, it is convenient to represent the S-modulus as a complex field S taking values in the upper half plane. An S-duality transformation

$$\gamma \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2;\mathbb{Z}) \tag{2.4}$$

acts simultaneously on the charges and the S-modulus by

$$\begin{pmatrix} N \\ M \end{pmatrix} \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} N \\ M \end{pmatrix}\,, \qquad S \to \frac{aS+b}{cS+d}\,. \tag{2.5}$$

To define T-duality transformations, it is convenient to represent the T-moduli by a 28×28 matrix μ_I^A satisfying

$$\mu^t\, L\, \mu = L \tag{2.6}$$

with the identification that $\mu \sim k\mu$ for every $k \in O(22;\mathbb{R}) \times O(6;\mathbb{R})$. Here L is the 28×28 matrix

$$L_{IJ} = \begin{pmatrix} -\mathbf{C}_{16} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathrm{I}_6 \\ \mathbf{0} & \mathrm{I}_6 & \mathbf{0} \end{pmatrix}, \tag{2.7}$$

with I_s the $s \times s$ identity matrix and $\mathbf{C}_{16}$ is the Cartan matrix of $E_8 \times E_8$. The T-moduli are then represented by the matrix

$$\mathcal{M} \;=\; \mu^t \mu \tag{2.8}$$

which satisifies

$$\mathcal{M}^t \;=\; \mathcal{M}\,, \qquad \mathcal{M}^t L \mathcal{M} = L\,. \tag{2.9}$$

In this basis, a T-duality transformation can then be represented by a 28×28 matrix R with integer entries satisfying

$$R^t L R \;=\; L\,, \tag{2.10}$$

which acts simultaneously on the charges and the T-moduli by

$$N \to RN; \quad M \to RM\,; \quad \mu \to \mu R^{-1} \tag{2.11}$$

Given the matrix μ_I^A, one obtains an embedding $\Lambda^{22,6} \subset \mathbb{R}^{22,6}$ of $\Pi^{22,6}$ which allows us to define the moduli-dependent charge vectors Q and P by

$$Q^A = \mu_I^A N_I\,, \qquad P^A = \mu_I^A M_I\,. \tag{2.12}$$

The matrix L has a 22-dimensional eigensubspace with eigenvalue -1 and a 6- dimensional eigensubspace with eigenvalue $+1$. Given Q and P, one can define the 'right-moving' and 'left-moving' charges[4] $Q_{R,L}$ and $P_{L,R}$ as the projections

$$Q_{R,L} \;=\; \frac{(1 \pm L)}{2} Q\,; \qquad P_{R,L} \;=\; \frac{(1 \pm L)}{2} P\,. \tag{2.13}$$

If the vectors N and M are nonparallel, then the state is quarter-BPS. On the other hand, if $N = pN_0$ and $M = qN_0$ for some $N_0 \in \Pi^{22,6}$ with p and q relatively prime integers, then the state is half-BPS.

An important piece of nonperturbative information about the dynamics of the theory is the exact spectrum of all possible dyonic BPS-states at all points in the moduli space. More specifically, one would like to compute the number $d(\Gamma)|_{S,\mu}$ of dyons of a given charge Γ at a specific point (S, μ) in the moduli space. Computation of these numbers is of course a very complicated dynamical problem. In fact, for a string compactification on a general Calabi-Yau threefold, the answer is not known. One main reason for focusing on this particular compactification on $K3 \times T^2$ is that in this case the dynamical problem has been essentially solved and the exact spectrum of dyons is now known. Furthermore, the results are easy to summarize and the numbers $d(\Gamma)|_{S,\mu}$ are given in terms of Fourier coefficients of various modular forms.

In view of the duality symmetries, it is useful to classify the inequivalent duality orbits labeled by various duality invariants. This leads to an interesting problem in number theory of classification of inequivalent duality orbits of various duality groups such as $SL(2,\mathbb{Z}) \times O(22,6;\mathbb{Z})$ in our case and more exotic groups like $E_{7,7}(\mathbb{Z})$ for other choices of compactification manifold X_6. It is important to remember though that a duality transformation acts simultaneously on charges and the moduli. Thus, it maps a state with charge Γ at a point in the moduli space (S, μ) to a state with charge Γ' but at some other point in the moduli space (S', μ'). In this respect, the half-BPS and quarter-BPS dyons behave differently.

---

[4] The right-moving charges couple to the graviphoton vector fields associated with the right-moving chiral currents in the conformal field theory of the dual heterotic string.

- For half-BPS states, the spectrum does not depend on the moduli. Hence $d(\Gamma)|_{S',\mu'} = d(\Gamma)|_{S,\mu}$. Furthermore, by an S-duality transformation one can choose a frame where the charges are purely electric with $M = 0$ and $N \neq 0$. Single-particle states have N primitive and the number of states depends only on the T-duality invariant integer $n \equiv N^2/2$. We can thus denote the degeneracy of half-BPS states $d(\Gamma)|_{S',\mu'}$ simply by $d(n)$.

- For quarter-BPS states, the spectrum does depend on the moduli, and $d(\Gamma)|_{S',\mu'} \neq d(\Gamma)|_{S,\mu}$. However, the partition function turns out to be independent of moduli and hence it is enough to classify the inequivalent duality orbits to label the partition functions. For the specific duality group $SL(2,\mathbb{Z}) \times O(22,6;\mathbb{Z})$ the partition functions are essentially labeled by a single discrete invariant [31, 5, 6].

  $$I \;=\; \gcd(N \wedge M)\,, \tag{2.14}$$

  The degeneracies themselves are Fourier coefficients of the partition function. For a given value of I, they depend only on[5] the moduli and the three T-duality invariants $(m, n, \ell) \equiv (M^2/2, N^2/2, N \cdot M)$. Integrality of (m, n, ℓ) follows from the fact that both N and M belong to $\Pi^{22,6}$. We can thus denote the degeneracy of these quarter-BPS states $d(\Gamma)|_{S,\mu}$ simply by $d(m, n, l)|_{S,\mu}$. For simplicity, we consider only $I = 1$ in this paper.

Given this classification, it is useful to choose a representative set of charges that can sample all possible values of the three T-duality invariants. For this purpose, we choose a point in the moduli space where the torus T^2 is a product of two circles $S^1 \times \tilde{S}^1$ and choose the following charges in a Type-IIB frame.

- For electric charges, we take n units of momentum along the circle S^1, and $\tilde{K}$ Kaluza-Klein monopoles associated with the circle $\tilde{S}^1$.
- For magnetic charges, we take Q_1 units of D1-brane charge wrapping S^1, Q_5 D5-brane wrapping $K3 \times S^1$ and l units of momentum along the $\tilde{S}^1$ circle.

We can thus write

$$\Gamma \;=\; \begin{bmatrix} N \\ M \end{bmatrix} \;=\; \begin{bmatrix} 0, & n; & 0, & \tilde{K} \\ Q_1, & \tilde{n}; & Q_5, & 0 \end{bmatrix}. \tag{2.15}$$

The T-duality quadratic invariants can be computed using a restriction of the matrix (2.7) to a $\Lambda^{(2,2)}$ Narain lattice of the form

$$L \;=\; \begin{pmatrix} \mathbf{0} & \mathrm{I}_2 \\ \mathrm{I}_2 & \mathbf{0} \end{pmatrix}, \tag{2.16}$$

---

[5] There is an additional dependence on arithmetic T-duality invariants but the degeneracies for states with nontrivial values of these T-duality invariants can be obtained from the degeneracies discussed here by demanding S-duality invariance [6].

to obtain

$$M^2/2 \; = \; Q_1 Q_5 \, , \quad N^2/2 \; = \; n\tilde{K} \, , \quad N \cdot M = \tilde{n}\tilde{K} \, . \tag{2.17}$$

We can simply the notation further by choosing $\tilde{K} = Q_5 = 1$, $Q_1 = m$, $\tilde{n} = l$ to obtain

$$M^2/2 \; = \; m, \quad N^2/2 \; = \; n, \quad N \cdot M = l \, . \tag{2.18}$$

For this set of charges, we can focus our attention on a subset of T-moduli associated with the torus T^2 parametrized by

$$\mathcal{M} \; = \; \begin{pmatrix} G^{-1} & G^{-1}B \\ -BG^{-1} & G - BG^{-1}B \end{pmatrix} \, , \tag{2.19}$$

where G_{ij} is the metric on the torus and B_{ij} is the antisymmetric tensor field. Let $U = U_1 + iU_2$ be the complex structure parameter, A be the area, and ϵ_{ij} be the Levi-Civita symbol with $\epsilon_{12} = -\epsilon_{21} = 1$, then

$$G_{ij} = \frac{A}{U_2} \begin{pmatrix} 1 & U_1 \\ U_1 & |U|^2 \end{pmatrix} \quad \text{and} \quad B_{ij} = AB\epsilon_{ij} \, , \tag{2.20}$$

and the complexified Kähler modulus $U = U_1 + iU_2$ is defined as $U := B + iA$. The S-modulus $S = S_1 + S_2$ is defined as

$$S := a + i \exp\left(-2\phi\right) \tag{2.21}$$

where a is the axion and ϕ is the dilaton field in the four dimensional heterotic frame. the relevant moduli can be parametrized by three complex scalars S, T, U which define the so-called 'STU' model in $\mathcal{N} = 2$ supergravity. Note that these moduli are labeled naturally in the heterotic frame which are related to the S_B, T_B, and U_B moduli in the Type-IIB frame by

$$S \; = \; U_B, \quad T = S_B, \quad U = T_B \, . \tag{2.22}$$

3. Modular forms in one variable

Before discussing mock modular forms, it is useful to recall the variety of modular objects that have already made their appearance in the context of counting black holes. In the following sections we give the basic definitions of modular forms, Jacobi forms, and Siegel forms, using the notations that are standard in the mathematics literature, and then in each case illustrate a physics application to counting quantum black holes by means of an example.

In the physics context, these modular forms arise as generating functions for counting various quantum black holes in string theory. The structure of poles of the counting function is of particular importance in physics, since it determines the asymptotic growth of the Fourier coefficients as well as the contour dependence of the Fourier coefficients which corresponds to the wall crossing phenomenon. These examples will also be relevant later in §11 in connection with mock modular forms. We suggest chapters I and III of [21] respectively as a good general reference for classical and Siegel modular forms and [57] for Jacobi modular forms.

3.1 Basic definitions and properties

Let $\mathbb{H}$ be the upper half plane, *i.e.*, the set of complex numbers τ whose imaginary part satisfies $\mathrm{Im}(\tau) > 0$. Let $SL(2,\mathbb{Z})$ be the group of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with integer entries such that $ad - bc = 1$.

A *modular form* $f(\tau)$ of weight k on $SL(2,\mathbb{Z})$ is a holomorphic function on $\mathbb{H}$, that transforms as

$$f(\frac{a\tau + b}{c\tau + d}) = (c\tau + d)^k \, f(\tau) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \,, \tag{3.1}$$

for an integer k (necessarily even if $f(0) \neq 0$). It follows from the definition that $f(\tau)$ is periodic under $\tau \to \tau + 1$ and can be written as a Fourier series

$$f(\tau) \;=\; \sum_{n=-\infty}^{\infty} a(n)\, q^n \qquad \big(q := e^{2\pi i \tau}\big)\,, \tag{3.2}$$

and is bounded as $\mathrm{Im}(\tau) \to \infty$. If $a(0) = 0$, then the modular form vanishes at infinity and is called a *cusp form.* Conversely, one may weaken the growth condition at ∞ to $f(\tau) = \mathrm{O}(q^{-N})$ rather than $\mathrm{O}(1)$ for some $N \geq 0$; then the Fourier coefficients of f have the behavior $a(n) = 0$ for $n < -N$. Such a function is called a *weakly holomorphic modular form.*

The vector space over $\mathbb{C}$ of holomorphic modular forms of weight k is usually denoted by M_k. Similarly, the space of cusp forms of weight k and the space of weakly holomorphic modular forms of weight k are denoted by S_k and $M_k^!$ respectively. We thus have the inclusion

$$S_k \;\subset M_k \;\subset\; M_k^! \,. \tag{3.3}$$

The Fourier coefficients of the modular forms in these spaces have different growth properties:

1. $f \in S_k \;\Rightarrow\; a_n \;=\; \mathrm{O}(n^{k/2})$ as $n \to \infty$;
2. $f \in M_k \;\Rightarrow\; a_n \;=\; \mathrm{O}(n^{k-1})$ as $n \to \infty$;
3. $f \in M_k^! \;\Rightarrow\; a_n \;=\; \mathrm{O}(e^{C\sqrt{n}})$ as $n \to \infty$ for some $C > 0$.

Some important modular forms on $SL(2,\mathbb{Z})$ are:

1. The *Eisenstein series* $E_k \in M_k$ $(k \geq 4)$. The first two of these are

$$E_4(\tau) \;=\; 1 \;+\; 240 \sum_{n=1}^{\infty} \frac{n^3 q^n}{1 - q^n} \;=\; 1 + 240q + 2160q^2 + \cdots \,, \tag{3.4}$$

$$E_6(\tau) \;=\; 1 \;-\; 504 \sum_{n=1}^{\infty} \frac{n^5 q^n}{1 - q^n} \;=\; 1 - 504q - 16632q^2 - \cdots \,. \tag{3.5}$$

2. The *discriminant function* Δ. It is given by the product expansion

$$\Delta(\tau) \;=\; q\prod_{n=1}^{\infty}(1-q^n)^{24} \;=\; q-24q^2+252q^3+... \tag{3.6}$$

or by the formula $\Delta = (E_4^3 - E_6^2)/1728$. We mention for later use that the function $E_2(\tau) = \dfrac{1}{2\pi i}\dfrac{\Delta'(\tau)}{\Delta(\tau)} = 1-24\sum\limits_{n=1}^{\infty}\dfrac{nq^n}{1-q^n}$ is also an Eisenstein series, but is not modular. (It is a so-called *quasimodular form*, meaning in this case that the non-holomorphic function $\widehat{E}_2(\tau) = E_2(\tau) - \dfrac{3}{\pi\,\mathrm{Im}(\tau)}$ transforms like a modular form of weight 2.) This function can be used to form the *Ramanujan-Serre derivative*

$$\partial_k^{\mathrm{RS}} \,:\, M_k \to M_{k+2}\,, \qquad \partial_k^{\mathrm{RS}} f(\tau) \,:=\, \frac{1}{2\pi i}\,f'(\tau) \,-\, \frac{k}{12}\,E_2(\tau)\,f(\tau)\,. \tag{3.7}$$

The ring of modular forms on $SL(2,\mathbb{Z})$ is generated freely by E_4 and E_6, so any modular form of weight k can be written (uniquely) as a sum of monomials $E_4^\alpha E_6^\beta$ with $4\alpha+6\beta = k$. We also have $M_k = \mathbb{C}\cdot E_k \oplus S_k$ and $S_k = \Delta\cdot M_{k-12}$, so any $f \in M_k$ also has a unique expansion as $\sum_{n=0}^{[k/12]}\alpha_n\,E_{k-12n}\,\Delta^n$ (with $E_0 = 1$ and E_2 replaced by 0). From either representation, we see that a modular form is uniquely determined by its weight and first few Fourier coefficients.

Given two modular forms (f,g) of weight (k,l), one can produce a sequence of modular forms of weight $k+l+2n$, $n \geq 0$ using the *Rankin-Cohen bracket*

$$[f,g]_n \;=\; [f,g]_n^{(k,l)} \;=\; \sum_{r+s=n}(-1)^s\binom{k+n-1}{r}\binom{\ell+n-1}{s}f^{(s)}(\tau)g^{(r)}(\tau) \tag{3.8}$$

where $f^{(m)} := \left(\frac{1}{2\pi i}\frac{d}{d\tau}\right)^m f$. For $n = 0$, this is simply the product of the two forms, while for $n > 0$ we always have $[f,g]_n \in S_{k+l+2n}$. The first two non-trivial examples are

$$[E_4,E_6]_1 \;=\; -3456\,\Delta\;, \qquad [E_4,E_4]_2 = 4800\,\Delta\;. \tag{3.9}$$

As well as modular forms on the full modular group $SL(2,\mathbb{Z})$, one can also consider modular forms on subgroups of finite index, with the same transformation law (3.1) and suitable conditions on the Fourier coefficients to define the notions of holomorphic, weakly holomorphic and cusp forms. The weight k now need no longer be even, but can be odd or even half integral, the easiest way to state the transformation property when $k \in \mathbb{Z}+\frac{1}{2}$ being to say that $f(\tau)/\theta(\tau)^{2k}$ is invariant under some congruence subgroup of $SL(2,\mathbb{Z})$, where $\theta(\tau) = \sum_{n\in\mathbb{Z}} q^{n^2}$. The graded vector space of modular forms on a fixed subgroup $\Gamma \subset SL(2,\mathbb{Z})$ is finite dimensional in each weight, finitely generated as an algebra, and closed under Rankin-Cohen brackets. Important examples of modular forms of half-integral weight are the *unary theta series*, i.e., theta series associated to a quadratic form in one variable. They come in two types:

$$\sum_{n\in\mathbb{Z}} \varepsilon(n)\, q^{\lambda n^2} \qquad \text{for some } \lambda \in \mathbb{Q}_+ \text{ and some even periodic function } \varepsilon \tag{3.10}$$

and

$$\sum_{n\in\mathbb{Z}} n\,\varepsilon(n)\, q^{\lambda n^2} \qquad \text{for some } \lambda \in \mathbb{Q}_+ \text{ and some odd periodic function } \varepsilon\,, \tag{3.11}$$

the former being a modular form of weight 1/2 and the latter a cusp form of weight 3/2. A theorem of Serre and Stark says that in fact every modular form of weight 1/2 is a linear combination of form of the type (3.10), a simple example being the identity

$$\eta(\tau) \;:=\; q^{1/24}\prod_{n=1}^{\infty}(1-q^n) \;=\; \sum_{n=1}^{\infty}\chi_{12}(n)\, q^{n^2/24}\,, \tag{3.12}$$

proved by Euler for the so-called *Dedekind eta function* $\eta(t) = \Delta(\tau)^{1/24}$. Here χ_{12} is the function of period 12 defined by

$$\chi_{12}(n) \;=\; \begin{cases} +1 \text{ if } n \equiv \pm 1 \,(\mathrm{mod}\, 12) \\ -1 \text{ if } n \equiv \pm 5 \,(\mathrm{mod}\, 12) \\ 0 \quad \text{if } (n, 12) > 1\ . \end{cases} \tag{3.13}$$

Finally, we recall the definition of the Petersson scalar product. If $f(\tau)$ and $g(\tau)$ are two modular forms of the same weight k and the same "multiplier system" on some subgroup Γ of finite index of $SL_2(\mathbb{Z})$ (this means that the quotient f/g is invariant under Γ), and if either $k < 1$ or else at least one of f and g is a cusp form, then we can define the (normalized) *Petersson scalar product* of f and g by

$$\big(f, g\big) \;=\; \int_{\mathcal{F}_\Gamma} f(\tau)\,\overline{g(\tau)}\,\tau_2^k\, d\mu(\tau) \Big/ \int_{\mathcal{F}_\Gamma} d\mu(\tau) \qquad \Big(d\mu(\tau) \;:=\; \frac{d\tau_1 d\tau_2}{\tau_2^2}\Big)\,, \tag{3.14}$$

where $\mathcal{F}_\Gamma$ is a fundamental domain for Γ. This definition is independent of the choice of the subgroup Γ and the fundamental domain $\mathcal{F}_\Gamma$. By the Rankin-Selberg formula (see, for example, [127]), we have that if $f = \sum_\lambda a_\lambda q^\lambda$ and $g = \sum_\lambda b_\lambda q^\lambda$ (where $\lambda \geq 0$ may have a denominator), then

$$(f, g) \;=\; \frac{\Gamma(k)}{(4\pi)^k}\, \mathrm{Res}_{s=k}\Bigg(\sum_{\lambda>0}\frac{a_\lambda\,\overline{b_\lambda}}{\lambda^s}\Bigg)\,, \tag{3.15}$$

a formula that will be used later. For instance, for $f = g = \eta$ we have $k = 1/2$ and $\sum_\lambda a_\lambda b_\lambda \lambda^{-s} = 24^s(1-2^{-2s})(1-3^{-2s})\,\zeta(2s)$, and hence $(\eta, \eta) = 1/\sqrt{6}$.

3.2 Quantum black holes and modular forms

Modular forms occur naturally in the context of counting the Dabholkar-Harvey (DH) states [39, 32], which are states in the string Hilbert space that are dual to perturbative BPS states. The spacetime helicity supertrace counting the degeneracies reduces to the partition function of a chiral conformal field theory on a genus-one worldsheet. The τ parameter above becomes the modular parameter of the genus one Riemann surface. The degeneracies are given by the Fourier coefficients of the partition function.

A well-known simple example is the partition function $Z(\tau)$ which counts the half-BPS DH states for the Type-II compactification on $K3 \times T^2$ considered here. In the notation of (2.3) these states have zero magnetic charge $M = 0$, but nonzero electric charge N with the T-duality invariant $N^2 = 2n$, which can be realized for example by setting $Q_1 = Q_5 = l = 0$ in (2.15). They are thus purely electric and perturbative in the heterotic frame[6]. The partition function is given by the partition function of the chiral conformal field theory of 24 left-moving transverse bosons of the heterotic string. The Hilbert space $\mathcal{H}$ of this theory is a unitary Fock space representation of the commutation algebra

$$[a_{in}, a^{\dagger}_{jm}] \;=\; \delta_{ij}\,\delta_{n+m,0} \qquad (i,\, j = 1, \ldots, 24\,, \quad n,\, m = 1,\, 2, \ldots, \infty) \tag{3.16}$$

of harmonic modes of oscillations of the string in 24 different directions. The Hamiltonian is

$$H \;=\; \sum_{i=1}^{24} n\, a^{\dagger}_{in}\, a_{in} \;-\; 1\,, \tag{3.17}$$

and the partition function is

$$Z(\tau) \;=\; \mathrm{Tr}_{\mathcal{H}}(q^{H})\,. \tag{3.18}$$

This can be readily evaluated since each oscillator mode of energy n contributes to the trace

$$1 + q^n + q^{2n} + \ldots \;=\; \frac{1}{1 - q^n}\,. \tag{3.19}$$

The partition function then becomes

$$Z(\tau) \;=\; \frac{1}{\Delta(\tau)}\,, \tag{3.20}$$

where Δ is the cusp form (3.6). Since Δ has a simple zero at $q = 0$, the partition function itself has a pole at $q = 0$, but has no other poles in $\mathbb{H}$. Hence, $Z(\tau)$ is a weakly holomorphic modular form of weight -12. This property is essential in the present physical context since it determines the asymptotic growth of the Fourier coefficients.

---

[6] Not all DH states are half-BPS. For example, the states that are perturbative in the Type-II frame correspond to a Type-II string winding and carrying momentum along a cycle in T^2. For such states both M and N are nonzero and nonparallel, and hence the state is quarter- BPS.

The degeneracy $d(n)$ of the state with electric charge N depends only on the T-duality invariant integer n and is given by

$$Z(\tau) \;=\; \sum_{n=-1}^{\infty} d(n)\, q^n \,. \tag{3.21}$$

For the Fourier integral

$$d(n) \;=\; \int_C e^{-2\pi i\tau n} Z(\tau) d\tau \,, \tag{3.22}$$

one can choose the contour $\mathcal{C}$ in $\mathbb{H}$ to be

$$0 \le \mathrm{Re}(\tau) < 1 \,, \tag{3.23}$$

for a fixed imaginary part $\mathrm{Im}(\tau)$. Since the partition function has no poles in $\mathbb{H}$ except at $q = 0$, smooth deformations of the contour do not change the Fourier coefficients and consequently the degeneracies $d(n)$ are uniquely determined from the partition function. This reflects the fact that the half-BPS states are immortal and do not decay anywhere in the moduli space. As a result, there is no wall crossing phenomenon, and no jumps in the degeneracy.

In number theory, the partition function above is well-known in the context of the problem of partitions of integers. We can therefore identify

$$d(n) \;=\; p_{24}(n+1) \qquad (n \ge 0) \,. \tag{3.24}$$

where $p_{24}(I)$ is the number of colored partitions of a positive integer I using integers of 24 different colors.

These states have a dual description in the Type-II frame where they can be viewed as bound states of Q_1 number of D1-branes and Q_5 number of D5-branes with $M^2/2 = Q_1 Q_5 \equiv m$. This corresponds to setting $n = \tilde{K} = l = 0$ in (2.15). In this description, the number of such bound states $d(m)$ equals the orbifold Euler character $\chi(\mathrm{Sym}^{m+1}(K3))$ of the symmetric product of $(m+1)$ copies of $K3$-surface [120]. The generating function for the orbifold Euler character

$$\hat{Z}(\sigma) = \sum_{m=-1}^{\infty} \chi(\mathrm{Sym}^{m+1}(K3))\, p^m \qquad \big(p := e^{2\pi i\sigma}\big) \tag{3.25}$$

can be evaluated [65] to obtain

$$\hat{Z}(\sigma) = \frac{1}{\Delta(\sigma)} \,. \tag{3.26}$$

Duality requires that the number of immortal BPS-states of a given charge must equal the number of BPS-states with the dual charge. The equality of the two partition functions (3.20) and (3.26) coming from two very different counting problems is consistent with this expectation. This fact was indeed one of the early indications of a possible duality between heterotic and Type-II strings [120].

The DH-states correspond to the microstates of a small black hole [103, 28, 40] for large n. The macroscopic entropy $S(n)$ of these black holes should equal the asymptotic growth of the degeneracy by the Boltzmann relation

$$S(n) = \log d(n); \quad n \gg 1\,. \tag{3.27}$$

In the present context, the macroscopic entropy can be evaluated from the supergravity solution of small black holes [81, 84, 83, 82, 28, 40]. The asymptotic growth of the microscopic degeneracy can be evaluated using the Hardy-Ramanujan expansion (Cardy formula). There is a beautiful agreement between the two results [28, 79]

$$S(n) \;=\; \log d(n) \sim 4\pi\sqrt{n} \quad n \gg 1\,. \tag{3.28}$$

Given the growth properties of the Fourier coefficients mentioned above, it is clear that, for a black hole whose entropy scales as a power of n and not as $\log(n)$, the partition function counting its microstates can be only weakly holomorphic and not holomorphic.

These considerations generalize in a straightforward way to congruence subgroups of $SL(2,\mathbb{Z})$ which are relevant for counting the DH-states in various orbifold compactifications with $\mathcal{N} = 4$ or $\mathcal{N} = 2$ supersymmetry [29, 104, 30].

4. Jacobi forms

4.1 Basic definitions

Consider a holomorphic function $\varphi(\tau, z)$ from $\mathbb{H} \times \mathbb{C}$ to $\mathbb{C}$ which is "modular in τ and elliptic in z" in the sense that it transforms under the modular group as

$$\varphi\Big(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\Big) \;=\; (c\tau+d)^k\, e^{\frac{2\pi i m c z^2}{c\tau+d}}\, \varphi(\tau, z) \qquad \forall \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2;\mathbb{Z}) \tag{4.1}$$

and under the translations of z by $\mathbb{Z}\tau + \mathbb{Z}$ as

$$\varphi(\tau, z+\lambda\tau+\mu) \;=\; e^{-2\pi i m(\lambda^2\tau+2\lambda z)}\varphi(\tau, z) \qquad \forall \quad \lambda,\, \mu \in \mathbb{Z}\,, \tag{4.2}$$

where k is an integer and m is a positive integer.

These equations include the periodicities $\varphi(\tau+1, z) = \varphi(\tau, z)$ and $\varphi(\tau, z+1) = \varphi(\tau, z)$, so φ has a Fourier expansion

$$\varphi(\tau, z) \;=\; \sum_{n,r} c(n, r)\, q^n\, y^r\,, \qquad\qquad (q := e^{2\pi i\tau},\; y := e^{2\pi i z})\;. \tag{4.3}$$

Equation (4.2) is then equivalent to the periodicity property

$$c(n,r) \;=\; C(4nm-r^2,\,r)\;, \qquad \text{where } C(\Delta,r) \text{ depends only on } r \pmod{2m}\;. \tag{4.4}$$

(We will sometimes denote $c(n,r)$ and $C(\Delta,r)$ by $c_\varphi(n,r)$ and $C_\varphi(\Delta,r)$ or by $c(\varphi;n,r)$ and $C(\varphi\,;\Delta,r)$ when this is required for emphasis or clarity.) The function $\varphi(\tau,z)$ is called a *holomorphic Jacobi form* (or simply a *Jacobi form*) of weight k and index m if the coefficients $C(\Delta,r)$ vanish for $\Delta<0$, *i.e.* if

$$c(n,r) \;=\; 0 \qquad \text{unless} \qquad 4mn \geq r^2\,. \tag{4.5}$$

It is called a *Jacobi cusp form* if it satisfies the stronger condition that $C(\Delta,r)$ vanishes unless Δ is strictly positive, *i.e.*

$$c(n,r)=0 \qquad \text{unless} \qquad 4mn > r^2\;, \tag{4.6}$$

and it is called a *weak Jacobi form* if it satisfies the weaker condition

$$c(n,r) \;=\; 0 \qquad \text{unless} \qquad n\geq 0 \tag{4.7}$$

rather than (4.5), whereas a merely *weakly holomorphic Jacobi form* satisfies only the yet weaker condition that $c(n,r)=0$ unless $n\geq n_0$ for some possibly negative integer n_0 (or equivalently $C(\Delta,r)=0$ unless $\Delta\geq\Delta_0$ for some possibly negative integer Δ_0). The space of all holomorphic (resp. cuspidal, weak, or weakly holomorphic) Jacobi forms of weight k and index m will be denoted by $J_{k,m}$ (resp. $J^0_{k,m}$, $\widetilde{J}_{k,m}$, or $\widetilde{J}^!_{k,m}$).

Finally, the quantity $\Delta=4mn-r^2$, which by virtue of the above discussion is the crucial invariant of a monomial q^ny^r occurring in the Fourier expansion of φ, will be referred to as its *discriminant*. (It would be mathematically more correct to use this word for the quantity $-\Delta$, but Δ is usually positive and it is more convenient to work with positive numbers.)

4.2 Theta expansion and Taylor expansion

A Jacobi form has two important representations, the *theta expansion* and the *Taylor expansion*. In this subsection, we explain both of these and the relation between them.

If $\varphi(\tau,z)$ is a Jacobi form, then the transformation property (4.2) implies its Fourier expansion with respect to z has the form

$$\varphi(\tau,z) \;=\; \sum_{\ell\,\in\,\mathbb{Z}} q^{\ell^2/4m}\; h_\ell(\tau)\; e^{2\pi i\ell z} \tag{4.8}$$

where $h_\ell(\tau)$ is periodic in ℓ with period $2m$. In terms of the coefficients (4.4) we have

$$h_\ell(\tau) \;=\; \sum_{\Delta} C(\Delta,\ell)\, q^{\Delta/4m} \qquad\qquad (\ell\,\in\,\mathbb{Z}/2m\mathbb{Z})\;. \tag{4.9}$$

Because of the periodicity property, equation (4.8) can be rewritten in the form

$$\varphi(\tau,z) = \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}} h_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z)\,, \tag{4.10}$$

where $\vartheta_{m,\ell}(\tau,z)$ denotes the standard index m theta function

$$\vartheta_{m,\ell}(\tau,z) \;:=\; \sum_{\substack{r\,\in\,\mathbb{Z}\\ r\,\equiv\,\ell\,(\mathrm{mod}\,2m)}} q^{r^2/4m}\,y^r \;\;=\; \sum_{n\,\in\,\mathbb{Z}} q^{(\ell+2mn)^2/4m}\,y^{\ell+2mn} \tag{4.11}$$

(which is a Jacobi form of weight $\frac{1}{2}$ and index m on some subgroup of $SL(2,\mathbb{Z})$). This is the theta expansion of φ. The coefficiens $h_\ell(\tau)$ are modular forms of weight $k-\frac{1}{2}$ and are weakly holomorphic, holomorphic or cuspidal if φ is a weak Jacobi form, a Jacobi form or a Jacobi cusp form, respectively. More precisely, the vector $h := (h_1,\ldots,h_{2m})$ transforms like a modular form of weight $k-\frac{1}{2}$ under $SL(2,\mathbb{Z})$.

The theta decomposition (4.10) leads to the definition of a differential operator on Jacobi forms as follows. Let

$$\mathcal{L}_m \;=\; \frac{4m}{2\pi i}\,\frac{\partial}{\partial\tau} \;-\; \frac{1}{(2\pi i)^2}\,\frac{\partial^2}{\partial z^2}$$

be the index m heat operator, which sends $\varphi = \sum c(n,r)q^ny^r$ to $\sum(4nm-r^2)c(n,r)q^ny^r$. Here the Fourier coefficients have the same periodicity property (4.4) as for φ, so $\mathcal{L}_m\varphi$ has the same elliptic transformation properties as a Jacobi form of index m. Moreover, since $\mathcal{L}_m$ annihilates all the $\vartheta_{m,\ell}$, we have $\mathcal{L}_m(\sum h_\ell\vartheta_{m,\ell}) = 4m\sum h_\ell'\vartheta_{m,\ell}$, so the modified heat operator

$$\mathcal{L}_{k,m} = \mathcal{L}_m - \frac{m(k-\frac{1}{2})}{3}\,E_2 \;\;:\;\; \sum_\ell h_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z) \;\mapsto\; 4m\sum_\ell \partial^{\mathrm{RS}}_{k-\frac{1}{2}} h_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z)\,, \tag{4.12}$$

where $\partial^{\mathrm{RS}}_{*}$ is the Ramanujan-Serre derivative defined in (3.7), sends $J_{k,m}$ to $J_{k+2,m}$ (and also $\widetilde{J}_{k,m}$ to $\widetilde{J}_{k+2,m}$ and $J^!_{k,m}$ to $J^!_{k+2,m}$). These operators will be used later.

A Jacobi form $\varphi \in \widetilde{J}_{k,m}$ also has a Taylor expansion in z which for k even takes the form

$$\varphi(\tau,z) \;=\; \xi_0(\tau) + \left(\frac{\xi_1(\tau)}{2} + \frac{m\xi_0'(\tau)}{k}\right)(2\pi iz)^2 + \left(\frac{\xi_2(\tau)}{24} + \frac{m\xi_1'(\tau)}{2\,(k+2)} + \frac{m^2\xi_0''(\tau)}{2k(k+1)}\right)(2\pi iz)^4 + \cdots \tag{4.13}$$

with $\xi_\nu \in M_{k+2\nu}(SL(2,\mathbb{Z}))$ and the prime denotes $\frac{1}{2\pi i}\frac{d}{d\tau}$ as before. In terms of the Fourier coefficients of φ, the modular form ξ_ν is given by

$$\frac{(k+2\nu-2)!}{(k+\nu-2)!}\,\xi_\nu(\tau) \;=\; \sum_{n=0}^{\infty}\left(\sum_r P_{\nu,k}(nm,r^2)c(n,r)\right)q^n\,, \tag{4.14}$$

where $P_{\nu,k}$ is a homogeneous polynomial of degree ν in r^2 and n with coefficients depending on k and m, the first few being

$$\begin{aligned} P_{0,k} &= 1 \,, \\ P_{1,k} &= kr^2 \,-\, 2nm \,, \\ P_{2,k} &= (k+1)(k+2)r^4 \,-\, 12(k+1)r^2mn \,+\, 12m^2n^2 \,. \end{aligned} \tag{4.15}$$

The Jacobi form φ is determined by the first $m+1$ coefficients ξ_ν, and the map $\varphi \mapsto (\xi_0, \dots, \xi_m)$ is an isomorphism from $\widetilde{J}_{k,m}$ to $M_k \oplus M_{k+2} \oplus \cdots \oplus M_{k+2m}$. For k odd, the story is similar except that (4.13) must be replaced by

$$\varphi(\tau, z) \;=\; \xi_0(\tau)\,(2\pi iz) \,+\, \left(\frac{\xi_1(\tau)}{6} + \frac{m\xi_0'(\tau)}{k+2}\right)(2\pi iz)^3 \,+\, \cdots \tag{4.16}$$

with $\xi_\nu \in M_{k+2\nu+1}(SL(2,\mathbb{Z}))$, and the map $\varphi \mapsto (\xi_0, \dots, \xi_{m-2})$ gives an isomorphism from $\widetilde{J}_{k,m}$ to $M_{k+1} \oplus M_{k+3} \oplus \cdots \oplus M_{k+2m-3}$.

We also observe that even if φ is weak, so that the individual coefficients $c(n,r)$ grow like $C^{\sqrt{4nm-r^2}}$, the coefficients $\sum_r P_{\nu,k}(nm, r^2)\, c(n,r)$ of ξ_ν still have only polynomial growth. We thus have the following descriptions (analogous to those given for classical modular forms in §3.1) of holomorphic and weak Jacobi forms in terms of their asymptotic properties:

Holomorphic Jacobi $\iff$ $c(n,r) = 0$ for $4mn - r^2 < 0$
$\iff$ the function $q^{m\alpha^2}\varphi(\tau, \alpha\tau+\beta)$ (which is a modular form of weight k and some level) is bounded as $\tau_2 \to \infty$ for every $\alpha, \beta \in \mathbb{Q}$
$\iff$ all $h_j(\tau)$ in (4.10) are bounded as $\tau_2 \to \infty$
$\iff$ $c(n,r)$ have polynomial growth.

Weak Jacobi $\iff$ $c(n,r) = 0$ for $n < 0$
$\iff$ $\varphi(\tau, \alpha\tau+\beta)$ is bounded as $\tau_2 \to \infty$ for any fixed $z \in \mathbb{C}$
$\iff$ all $h_j(\tau) = O(q^{-j^2/4m})$ as $\tau_2 \to \infty$
$\iff$ $\sum_r P_{\nu,k}(nm, r^2)c(n,r)$ have polynomial growth.

Finally, the relation between the Taylor expansion (4.13) of a Jacobi form and its theta expansion (4.10) is given by

$$\xi_\nu(\tau) \;=\; (4m)^\nu \binom{k+2\nu-2}{\nu}^{-1} \sum_{\ell\,(\mathrm{mod}\,2m)} \big[h_\ell(\tau),\, \vartheta^0_{m,\ell}(\tau)\big]_\nu \;, \tag{4.17}$$

where $[\ ,\ ]_\nu = [\ ,\ ]_\nu^{(k-\frac{1}{2},\frac{1}{2})}$ denotes the Rankin-Cohen bracket (which, as we mentioned above, also works in half-integral weight), and $\vartheta^0_{m,\ell}(\tau) = \vartheta_{m,\ell}(\tau, 0)$ (Thetanullwerte). There is a similar formula in the odd case, but with $\vartheta^0_{m,\ell}(\tau)$ replaced by

$$\vartheta^1_{m,\ell}(\tau) \;=\; \frac{1}{2\pi i}\,\frac{\partial}{\partial z}\vartheta_{m,\ell}(\tau,z)\Big|_{z=0} \;=\; \sum_{r\equiv\ell\,(\mathrm{mod}\,2m)} r\,q^{r^2/4m}\,. \tag{4.18}$$

4.3 Example: Jacobi forms of index 1

If $m=1$, (4.4) reduces to $c(n,r)=C(4n-r^2)$ where $C(\Delta)$ is a function of a single argument, because $4n-r^2$ determines the value of $r\,(\mathrm{mod}\,2)$. So any $\varphi\in J^{\mathrm{weak}}_{k,1}$ has an expansion of the form

$$\varphi(\tau,z) \;=\; \sum_{n,r\in\mathbb{Z}} C(4n-r^2)\,q^n\,y^r\,. \tag{4.19}$$

It also follows that k must be even, since in general, $C(\Delta,-r)=(-1)^kC(\Delta,r)$.

One has the isomorphisms $J_{k,1}\cong M_k\oplus S_{k+2}$ and $J^{\mathrm{weak}}_{k,1}\cong M_k\oplus M_{k+2}$. If $\varphi\in J^{\mathrm{weak}}_{k,1}$ with an expansion as in (4.19), then

$$\varphi(\tau,0) \;=\; \sum_{n=0}^{\infty} a(n)\,q^n\,, \qquad \frac{1}{2(2\pi i)^2}\varphi''(\tau,0) \;=\; \sum_{n=0}^{\infty} b(n)\,q^n\,, \tag{4.20}$$

where

$$a(n) \;=\; \sum_{r\in\mathbb{Z}} C(4n-r^2)\,, \qquad b(n) \;=\; \sum_{r>0} r^2\,C(4n-r^2)\,, \tag{4.21}$$

and the isomorphisms are given (if $k>0$) by the map $\varphi\mapsto(\mathcal{A},\mathcal{B})$ with

$$\mathcal{A}(\tau) \;=\; \sum a(n)\,q^n\in M_k\,, \qquad \mathcal{B}(\tau) \;=\; \sum\bigl(kb(n)-na(n)\bigr)\,q^n\in M_{k+2}\,. \tag{4.22}$$

For $J_{k,1}$ one also has the isomorphism $J_{k,1}\cong M^+_{k-\frac{1}{2}}(\Gamma_0(4))$ given by

$$\varphi(\tau,z)\leftrightarrow g(\tau) \;=\; \sum_{\substack{\Delta\geq 0\\ \Delta\equiv 0,\,3\,\mathrm{mod}4}} C(\Delta)\,q^{\Delta}\,. \tag{4.23}$$

We have four particularly interesting examples $\varphi_{k,1}$

$$\varphi_{k,1}(\tau,z) \;=\; \sum_{n,\,r\in\mathbb{Z}} C_k(4n-r^2)\,q^n\,y^r\,, \quad k=-2,\,0,\,10,\,12\,, \tag{4.24}$$

which have the properties (defining them uniquely up to multiplication by scalars)

- $\varphi_{10,1}$ and $\varphi_{12,1}$ are the two index 1 Jacobi cusp forms of smallest weight;
- $\varphi_{-2,1}$ and $\varphi_{0,1}$ are the unique weak Jacobi forms of index 1 and weight ≤ 0;

- $\varphi_{-2,1}$ and $\varphi_{0,1}$ generate the ring of weak Jacobi forms of even weight freely over the ring of modular forms of level 1, so that

$$J_{k,m}^{\text{weak}} \;=\; \bigoplus_{j=0}^{m} M_{k+2j}\left(SL(2,\mathbb{Z})\right)\cdot\varphi_{-2,1}^{j}\,\varphi_{0,1}^{m-j} \qquad (k \text{ even})\,; \tag{4.25}$$

- $\varphi_{-2,1} = \varphi_{10,1}/\Delta$, $\varphi_{0,1} = \varphi_{12,1}/\Delta$, and the quotient $\varphi_{0,1}/\varphi_{-2,1} = \varphi_{12,1}/\varphi_{10,1}$ is a multiple of the Weierstrass $\wp$ function.

The Fourier coefficients of these functions can be computed from the above recursions, since the pairs $(\mathcal{A},\mathcal{B})$ for $\varphi = \varphi_{-2,1}$, $\varphi_{0,1}$, $\varphi_{10,1}$ and $\varphi_{12,1}$ are proportional to $(0,1)$, $(1,0)$, $(0,\Delta)$ and $(\Delta,0)$, respectively[7]. The results for the first few Fourier coefficients are given in Table 4.3 below. In particular, the Fourier expansions of $\varphi_{-2,1}$ and $\varphi_{0,1}$ begin

$$\varphi_{-2,1} \;=\; \frac{(y-1)^2}{y} \;-\; 2\,\frac{(y-1)^4}{y^2}\,q \;+\; \frac{(y-1)^4(y^2-8y+1)}{y^3}\,q^2 \;+\;\cdots\;, \tag{4.26}$$

$$\varphi_{0,1} \;=\; \frac{y^2+10y+1}{y} \;+\; 2\frac{(y-1)^2\,(5y^2-22y+5)}{y^2}\,q \;+\;\cdots\;. \tag{4.27}$$

k	$C_k(-1)$	$C_k(0)$	$C_k(3)$	$C_k(4)$	$C_k(7)$	$C_k(8)$	$C_k(11)$	$C_k(12)$	$C_k(15)$
-2	1	-2	8	-12	39	-56	152	-208	513
0	1	10	-64	108	-513	808	-2752	4016	-11775
10	0	0	1	-2	-16	36	99	-272	-240
12	0	0	1	10	-88	-132	1275	736	-8040

Table 1: Some Fourier coefficients

The functions $\varphi_{k,1}$ ($k = 10,\ 0,\ -2$) can be expressed in terms of the Dedekind eta function (3.12) and the Jacobi theta functions ϑ_1, ϑ_2, ϑ_3, ϑ_4 by the formlas

$$\varphi_{10,1}(\tau,z) \;=\; \eta^{18}(\tau)\,\vartheta_1^2(\tau,z)\,, \tag{4.28}$$

$$\varphi_{-2,1}(\tau,z) \;=\; \frac{\vartheta_1^2(\tau,z)}{\eta^6(\tau)} \;=\; \frac{\varphi_{10,1}(\tau,z)}{\Delta(\tau)}\;. \tag{4.29}$$

$$\varphi_{0,1}(\tau,z) \;=\; 4\left(\frac{\vartheta_2(\tau,z)^2}{\vartheta_2(\tau)^2}+\frac{\vartheta_3(\tau,z)^2}{\vartheta_3(\tau)^2}+\frac{\vartheta_4(\tau,z)^2}{\vartheta_4(\tau)^2}\right)\,, \tag{4.30}$$

Finally, we say a few words about Jacobi forms of odd weight. Such a form cannot have index 1, as we saw. In index 2, the isomorphisms $J_{k,2} \cong S_{k+1}$ and $J_{k,2}^{\text{weak}} \cong M_{k+1}$ show that the first examples of holomorphic and weak Jacobi forms occur in weights 11 and −1, respectively, and are related by $\varphi_{-1,2} = \varphi_{11,2}/\Delta$. The function $\varphi_{-1,2}$ is given explicitly by

---

[7] For $k = 0$, the second formula in (4.22) must be modified, and the function $\sum b(n)q^n$ for $\varphi_{0,1}$ is in fact $E_2(\tau)$.

$$\varphi_{-1,2}(\tau, z) \;=\; \frac{\vartheta_1(\tau, 2z)}{\eta^3(\tau)}\;, \tag{4.31}$$

with Fourier expansion beginning

$$\varphi_{-1,2} \;=\; \frac{y^2-1}{y} \;-\; \frac{(y^2-1)^3}{y^3}\,q \;-\; 3\,\frac{(y^2-1)^3}{y^3}\,q^2 \;+\; \cdots\;, \tag{4.32}$$

and its square is related to the index 1 Jacobi forms defined above by

$$432\,\varphi_{-1,2}^2 \;=\; \varphi_{-2,1}\left(\varphi_{0,1}^3 \;-\; 3\,E_4\,\varphi_{-2,1}^2\,\varphi_{0,1} \;+\; 2\,E_6\,\varphi_{-2,1}^3\right)\,. \tag{4.33}$$

(In fact, $\varphi_{-1,2}/\varphi_{-2,1}^2$ is a multiple of $\wp'(\tau,z)$ and this equation, divided by $\varphi_{-2,1}^4$, is just the usual equation expressing $\wp'^{\,2}$ as a cubic polynomial in $\wp$.) It is convenient to introduce abbreviations

$$A \;=\; \varphi_{-2,1}\,, \qquad B \;=\; \varphi_{0,1}\,, \qquad C \;=\; \varphi_{-1,2}\,. \tag{4.34}$$

With these notations, the structure of the full bigraded ring of weak Jacobi forms is given by

$$J_{*,*}^{\text{weak}} \;=\; \mathbb{C}[E_4, E_6, A, B, C]\,\big/\,(432C^2 - AB^3 + 3E_4A^3B - 2E_6A^4)\,. \tag{4.35}$$

4.4 Hecke-like operators

In [57] Hecke operators T_ℓ acting on $J_{k,m}$ were introduced, but also various "Hecke-like" operators, again defined by the action of certain combinations of elements of $GL(2,\mathbb{Q}) \rtimes \mathbb{Q}^2$, which send Jacobi forms to Jacobi forms, but now possibly changing the index. We describe these operators here. This subsection is more technical than the preceding three, but will be essential later.

The first is the very simple operator U_s $(s \geq 1)$ which sends $\varphi(\tau, z)$ to $\varphi(\tau, sz)$, i.e.,

$$U_s\;:\quad \sum_{n,r} c(n,r)\,q^n\,y^r \quad\mapsto\quad \sum_{n,r} c(n,r)\,q^n\,y^{sr}\;, \tag{4.36}$$

This operator maps $J_{k,m}$ to J_{k,s^2m}.

The second operator, $V_{k,t}$ $(t \geq 1)$, sends $J_{k,m}$ to $J_{k,tm}$. It is given in terms of its action on Fourier coefficients by

$$c(\varphi|V_{k,t}\,;n,r) \;=\; \sum_{d|(n,r,t)} d^{k-1}c\Big(\varphi\,;\frac{nt}{d^2},\frac{r}{d}\Big)\,, \qquad C(\varphi|V_{k,t}\,;\Delta,r) \;=\; \sum_{d|(\frac{r^2+\Delta}{4mt},r,t)} d^{k-1}C\Big(\varphi\,;\frac{\Delta}{d^2},\frac{r}{d}\Big)\,. \tag{4.37}$$

In §9, we will also need the modified operator $\mathcal{V}_{k,t}^{(m)} : J_{k,m} \mapsto J_{k,tm}$ defined by

$$\mathcal{V}_{k,t}^{(m)} \;=\; \sum_{\substack{s^2|t\\(s,m)=1}} \mu(s)\, s^{k-1}\, V_{k,t/s^2}\, U_s\,, \tag{4.38}$$

so that, for instance, $\mathcal{V}_{k,t}^{(m)} = V_{k,t}$ if t is square-free, but $\mathcal{V}_{k,4}^{(m)} = V_{k,4} - 2^{k-1}U_2$ if m is odd.

Next, for each positive integer m_1 with $m_1 \| m$ (this means that $m = m_1 m_2$ with m_1 and m_2 coprime) we have an involution W_{m_1} on $J_{k,m}$, which we will call an *Atkin-Lehner involution,*[8] that is defined in terms of the theta expansion of Jacobi forms by

$$W_{m_1}\;:\quad \sum_{\ell\;(\mathrm{mod}\,2m)} h_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z) \quad\mapsto\quad \sum_{\ell\;(\mathrm{mod}\,2m)} h_{\ell^*}(\tau)\,\vartheta_{m,\ell}(\tau,z) \tag{4.39}$$

(or equivalently by $C(\Delta,r) \mapsto C(\Delta,r^*)$), where the involution $\ell \mapsto \ell^*$ on $\mathbb{Z}/2m\mathbb{Z}$ is defined by

$$\ell^* \equiv -\ell \,(\mathrm{mod}\,2m_1), \qquad \ell^* \equiv +\ell\,(\mathrm{mod}\,2m_2)\;. \tag{4.40}$$

These operators commute and satisfy $W_{m/m_1} = (-1)^k W_{m_1}$, so that we get an eigenspace decomposition

$$J_{k,m} \;=\; \bigoplus_{\substack{\boldsymbol{\varepsilon}=(\varepsilon_1,\ldots,\varepsilon_t)\,\in\,\{\pm1\}^t\\ \varepsilon_1\cdots\varepsilon_t\,=\,(-1)^k}} J_{k,m}^{\boldsymbol{\varepsilon}}\;, \tag{4.41}$$

where $m = p_1^{r_1}\cdots p_t^{r_t}$ is the prime power decomposition of m and ε_i is the eigenvalue of $W_{p_i^{r_i}}$. One can also introduce the "p-adic Atkin-Lehner involution" W_{p^∞}, which acts on any $J_{k,m}$ as W_{p^ν} where $p^\nu \| m$. (In particular, it acts as the identity if $p \nmid m$.) This operator simply changes the second index r in $C(\Delta,r)$ to r^*, where r is equal to the negative of r^* p-adically and to r^* p'-adically for any $p' \neq p$. In particular, W_{p^∞} commutes with all U_s, $V_{k,t}$, and $\mathcal{V}_{k,t}^m$.

We now describe a further Hecke-like operator, which was not given in [57] but is mentioned briefly in [115], p. 139, and which we will also need in the sequel. If t is a positive integer whose square divides m, then we define $u_t:\; J_{k,m} \;\longrightarrow J_{k,m/t^2}$ by

$$u_t\;:\quad \sum_{n,r} c(n,r)\,q^n\,y^r \quad\mapsto\quad \sum_{n,r}\Bigg(\sum_{a\;(\mathrm{mod}\;t)} c\big(n + ar + \frac{m}{t^2}a^2,\; tr + \frac{2ma}{t}\big)\Bigg)\,q^n\,y^r \tag{4.42}$$

or equivalently by

$$C_{\varphi|u_t}(\Delta,\;\ell\;(\mathrm{mod}\;2m/t^2)) \;=\; \sum_{\substack{r\;(\mathrm{mod}\;2m/t)\\ r\equiv\ell\;(\mathrm{mod}\;2m/t^2)}} C_\varphi(\Delta t^2,\; rt\;(\mathrm{mod}\;2m))\;. \tag{4.43}$$

---

[8] This terminology was not used in [57], where these involutions were introduced, but is justified by the results of [115] where it is shown that they correspond in a precise sense to the Atkin-Lehner involutions on the space of modular forms of weight $2k-2$ and level m.

It is easily checked that this map does indeed send Jacobi forms to Jacobi forms of the index stated, and that for any integers m, $t > 0$, the composite map $J_{k,m/t^2} \xrightarrow{U_t} J_{k,m} \xrightarrow{u_t} J_{k,m/t^2}$ is simply multiplication by t (because, if $\varphi = \varphi_1 \,|\, U_t$ with φ_1 of index m/t^2, then each term on the right of (4.43) equals $C_{\varphi_1}(\Delta, \ell)$). It follows that we have

$$J_{k,m} \;=\; \bigoplus_{t^2|m} J^{\text{prim}}_{k,\frac{m}{t^2}} \,|\, U_t \,, \tag{4.44}$$

where

$$J^{\text{prim}}_{k,m} \;=\; \bigcap_{\substack{t^2|m\\ t>1}} \ker\big(J_{k,m} \xrightarrow{u_t} J_{k,m/t^2}\big) \;=\; \bigcap_{\substack{p^2|m\\ p\text{ prime}}} \ker\big(J_{k,m} \xrightarrow{u_p} J_{k,m/p^2}\big) \tag{4.45}$$

(the equivalence of the two definitions follows because $u_t\, u_{t'} \,=\, u_{tt'}$ for any t, t'), with the projection map $\pi^{\text{prim}} : J_{k,m} \mapsto J^{\text{prim}}_{k,m}$ onto the first factor in (4.44) given by

$$\pi^{\text{prim}}(\varphi) \;=\; \sum_{t^2|m} \frac{\mu(t)}{t}\, \varphi \,|\, u_t \,|\, U_t \,. \tag{4.46}$$

We will sometimes find it useful to work, not with the space of Jacobi forms, but with the larger space of holomorphic functions on $\mathbb{H} \times \mathbb{C}$ that obey the elliptic transformation law (4.2) of Jacobi forms, but are not required to obey the modular transformation property (4.1). We shall call such functions *elliptic forms*. (The word *Jacobi-like forms* has been used in the literature to denote functions which, conversely, satisfy (4.1) but not (4.2).) Any elliptic form has a Fourier expansion as in (4.3) with the periodicity condition (4.4). We will denote by $\mathcal{E}_m$, $\mathcal{E}^0_m$, $\widetilde{\mathcal{E}}_m$ the spaces of all elliptic forms of index m satisfying the growth conditions (4.5), (4.6), and (4.7), respectively, and we will call them *holomorphic, cuspidal*, and *weak* respectively. We also denote by $\mathcal{E}_{\pm,m}$, $\mathcal{E}^0_{\pm,m}$, $\widetilde{\mathcal{E}}_{\pm,m}$ the $(\pm1)-$eigenspaces of theta functions under the involution $\varphi(\tau, z) \mapsto \varphi(\tau, -z)$, i.e., those functions whose Fourier coefficients $c(n, r)$ satisfy the parity condition $c(n, -r) = \pm c(n, r)$. Because any Jacobi form of weight k satisfies this parity condition with $\pm1 = (-1)^k$ (by (4.1) applied to $-\mathbf{1}_2$), we have

$$J_{k,m} \subset \mathcal{E}_{\pm,m}\,, \quad J^0_{k,m} \subset \mathcal{E}^0_{\pm,m}\,, \quad \widetilde{J}_{k,m} \subset \widetilde{\mathcal{E}}_{\pm,m} \qquad \big((-1)^k = \pm1\big)\,. \tag{4.47}$$

It is now clear from the definitions above that all of the Hecke-like operators V_t, U_t, W_{m_1} for $(m_1, m/m_1) = 1$ and u_t for $t^2|m$ defined in this section extend from the space of Jacobi forms to the space of elliptic forms, and that the decompositions (4.41) and (4.44) remain true if we replace all the J_ks by $\mathcal{E}_+$s.

Finally, we should warn the reader explicitly that, although the operators U_t and V_t also act on the spaces $\widetilde{J}_{k,*}$ and $\widetilde{\mathcal{E}}_{\pm,*}$ of weak Jacobi or weak elliptic forms, the operators W_{m_1} and u_t do not: the condition (4.7) is equivalent to saying that $C(\Delta, \ell)$ for $|\ell| \le m$ vanishes whenever $\Delta <$

$-\ell^2$, and this condition is preserved neither under the permutation $\ell \pmod{2m} \mapsto \ell^* \pmod{2m}$ defining W_{m_1}, nor by the summation conditions in (4.43). As a concrete example, the 7-dimensional space $\widetilde{J}_{2,6}$ contains a 6-dimensional subspace of forms invariant under the involution W_2, but no anti-invariant form, and its image under the projection operator $\pi_2^- = \frac{1}{2}(1 - W_2)$ to anti-invariant forms is the 1-dimensional space spanned by

$$\varphi_{2,6}^-(\tau, z) \;=\; j'(\tau)\,\frac{\vartheta_1(\tau, 4z)}{\vartheta_1(\tau, 2z)} \;=\; \sum_{n,\,r\in\mathbb{Z}} \chi_{12}(r)\, C_{2,6}(24n - r^2)\; q^n\, y^r\,, \tag{4.48}$$

with

Δ	≤ -49	-25	-1	23	47	71	$\cdots$
$C_{2,6}(\Delta)$	0	-1	-1	196882	43184401	2636281193	$\cdots$

,

which is a weakly holomorphic but not a weak Jacobi form. (Specifically, for any $\varphi \in \widetilde{J}_{2,6}$ we have $\pi_2^-(\varphi) = 6\, c_\varphi(0,5)\, \varphi_{2,6}^-\,$.)

4.5 Quantum black holes and Jacobi forms

Jacobi forms usually arise in string theory as elliptic genera of two-dimensional superconformal field theories (SCFT) with $(2,2)$ or more worldsheet supersymmetry[9]. We denote the superconformal field theory by $\sigma(\mathcal{M})$ when it corresponds to a sigma model with a target manifold $\mathcal{M}$. Let H be the Hamiltonian in the Ramond sector, and J be the left-moving $U(1)$ R-charge. The elliptic genus $\chi(\tau, z; \mathcal{M})$ is then defined as [123, 3, 97] a trace over the Hilbert space $\mathcal{H}_R$ in the Ramond sector

$$\chi(\tau, z; \mathcal{M}) \;=\; \mathrm{Tr}_{\mathcal{H}_R}\big((-1)^F q^H y^J\big)\,, \tag{4.49}$$

where F is the fermion number.

An elliptic genus so defined satisfies the modular transformation property (4.1) as a consequence of modular invariance of the path integral. Similarly, it satisfies the elliptic transformation property (4.2) as a consequence of spectral flow. Furthermore, in a unitary SCFT, the positivity of the Hamiltonian implies that the elliptic genus is a weak Jacobi form. The decomposition (4.10) follows from bosonizing the $U(1)$ current in the standard way so that the contribution to the trace from the momentum modes of the boson can be separated into the theta function (4.11). See, for example, [77, 93] for a discussion. This notion of the elliptic genus can be generalized to a $(0,2)$ theory using a left-moving $U(1)$ charge J which may not be an R- charge. In this case spectral flow is imposed as an additional constraint and follows from gauge invariance under large gauge transformations [44, 61, 80, 48].

A particularly useful example in the present context is $\sigma(K3)$, which is a $(4,4)$ SCFT whose target space is a $K3$ surface. The elliptic genus is a topological invariant and is independent of the moduli of the $K3$. Hence, it can be computed at some convenient point in the $K3$ moduli space, for example, at the orbifold point where the $K3$ is the Kummer surface. At this point, the $\sigma(K3)$ SCFT can be regarded as a $\mathbb{Z}_2$ orbifold of the $\sigma(T^4)$ SCFT, which is an SCFT with a torus T^4 as the target space. A simple computation using standard techniques of orbifold conformal field theory [43, 63] yields

---

[9] An SCFT with (r,s) supersymmetries has r left-moving and s right-moving supersymmetries.

$$\chi(\tau, z; K3) \;=\; 2\,\varphi_{0,1}(\tau, z) \;=\; 2\sum C_0(4n - l^2)\, q^n\, y^l\,. \tag{4.50}$$

Note that for $z = 0$, the trace (4.49) reduces to the Witten index of the SCFT and correspondingly the elliptic genus reduces to the Euler character of the target space manifold. In our case, one can readily verify from (4.50) and (4.30) that $\chi(\tau, 0; K3)$ equals 24, which is the Euler character of $K3$.

A well-known physical application of Jacobi forms is in the context of the five-dimensional Strominger-Vafa black hole[118], which is a bound state of Q_1 D1-branes, Q_5 D5-branes, n units of momentum and l units of five-dimensional angular momentum [12]. The degeneracies $d_m(n, l)$ of such black holes depend only on $m = Q_1Q_5$. They are given by the Fourier coefficients $c(n, l)$ of the elliptic genus $\chi(\tau, z; \mathrm{Sym}^{m+1}(K3))$ of symmetric product of $(m+1)$ copies of $K3$-surface.

Let us denote the generating function for the elliptic genera of symmetric products of K3 by

$$\widehat{Z}(\sigma, \tau, z) := \sum_{m=-1}^{\infty} \chi(\tau, z; \mathrm{Sym}^{m+1}(K3))\, p^m \tag{4.51}$$

where $\chi_m(\tau, z)$ is the elliptic genus of $\mathrm{Sym}^m(K3)$. A standard orbifold computation [50] gives

$$\widehat{Z}(\sigma, \tau, z) = \frac{1}{p} \prod_{r>0,\; s\geq 0,\; t} \frac{1}{(1 - q^s y^t p^r)^{2C_0(rs,t)}} \tag{4.52}$$

in terms of the Fourier coefficients $2C_o$ of the elliptic genus of a single copy of $K3$.

For $z = 0$, it can be checked that, as expected, the generating function (4.52) for elliptic genera reduces to the generating function (3.26) for Euler characters

$$\widehat{Z}(\sigma, \tau, 0) = \hat{Z}(\sigma) = \frac{1}{\Delta(\sigma)}\,. \tag{4.53}$$

5. Siegel modular forms

5.1 Definitions and examples of Siegel modular forms

Let $Sp(2, \mathbb{Z})$ be the group of (4×4) matrices g with integer entries satisfying $gJg^t = J$ where

$$J \equiv \begin{pmatrix} 0 & -I_2 \\ I_2 & 0 \end{pmatrix} \tag{5.1}$$

is the symplectic form. We can write the element g in block form as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{5.2}$$

where A, B, C, D are all (2×2) matrices with integer entries. Then the condition $gJg^t = J$ implies

$$AB^t = BA^t, \qquad CD^t = DC^t, \qquad AD^t - BC^t = \mathbf{1}\,, \tag{5.3}$$

Let $\mathbb{H}_2$ be the (genus two) Siegel upper half plane, defined as the set of (2×2) symmetric matrix Ω with complex entries

$$\Omega = \begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} \tag{5.4}$$

satisfying

$$\mathrm{Im}(\tau) > 0, \quad \mathrm{Im}(\sigma) > 0, \quad \det(\mathrm{Im}(\Omega)) > 0\,. \tag{5.5}$$

An element $g \in Sp(2,\mathbb{Z})$ of the form (5.2) has a natural action on $\mathbb{H}_2$ under which it is stable:

$$\Omega \to (A\Omega + B)(C\Omega + D)^{-1}. \tag{5.6}$$

The matrix Ω can be thought of as the period matrix of a genus two Riemann surface[10] on which there is a natural symplectic action of $Sp(2,\mathbb{Z})$.

A Siegel form $F(\Omega)$ of weight k is a holomorphic function $\mathbb{H}_2 \to \mathbb{C}$ satisfying

$$F\big((A\Omega + B)(C\Omega + D)^{-1}\big) \;=\; \det(C\Omega + D)^k\, F(\Omega)\,. \tag{5.7}$$

A Siegel modular form can be written in terms of its Fourier series

$$F(\Omega) \;=\; \sum_{\substack{n,\,r,\,m\,\in\,\mathbb{Z} \\ r^2 \le 4mn}} a(n,r,m)\, q^n\, y^r\, p^m\,. \tag{5.8}$$

If one writes this as

$$F(\Omega) = \sum_{m=0}^{\infty} \varphi^{\mathrm{F}}_m(\tau,z)\, p^m \tag{5.9}$$

with

$$\varphi^F_m(\tau,z) \;=\; \sum_{n,\,r} a(n,r,m)\, q^n\, y^r\,, \tag{5.10}$$

then each $\varphi^{\mathrm{F}}_m (m \ge 0)$ is a Jacobi form of weight k and index m.

---

[10] See [59, 41, 9] for a discussion of the connection with genus-two Riemann surfaces.

An important special class of Siegel forms were studied by Maass which he called the *Spezialschar*. They have the property that $a(n,r,m)$ depends only on the discriminant $4mn-r^2$ if (n,r,m) are coprime, and more generally

$$a(n,r,m) \;=\; \sum_{d|(n,r,m),\; d>0} d^{k-1} C\Big(\frac{4mn-r^2}{d^2}\Big) \tag{5.11}$$

for some coefficients $C(N)$. Specializing to $m=1$, we can see that these numbers are simply the coefficients associated via (4.19) to the Jacobi form $\varphi = \varphi_1^{\mathrm{F}} \in J_{k,1}$, and that (5.11) says precisely that the other Fourier-Jacobi coefficients of F are given by $\varphi_m^{\mathrm{F}} = \varphi_1^{\mathrm{F}}|V_m$ with V_m as in (4.37). Conversely, if φ is any Jacobi form of weight k and index 1 with Fourier expansion (4.19), then the function $F(\Omega)$ defined by (5.8) and (5.11) or by $F(\Omega) = \sum_{m=0}^{\infty}\big(\varphi|V_m\big)(\tau,z)\,p^m$ is a Siegel modular form of weight k with $\varphi_1^{\mathrm{F}} = \varphi$. The resulting map from $J_{k,1}$ to the Spezialschar is called the Saito-Kurokawa lift or *additive lift* since it naturally gives the sum representation of a Siegel form using the Fourier coefficients of a Jacobi form as the input. (More information about the additive lift can be found in [57].)

The example of interest to us is the Igusa cusp form Φ_{10} (the unique cusp form of weight 10) which is the Saito-Kurokawa lift of the Jacobi form $\varphi_{10,1}$ introduced earlier, so that

$$\Phi_{10}(\Omega) \;=\; \sum_{m=1}^{\infty}\big(\varphi_{10,1}|V_m\big)(\tau,z)\,p^m \;=\; \sum_{n,\,r,\,m} a_{10}(n,r,m)\,q^n\,y^r\,p^m\,, \tag{5.12}$$

where a_{10} is defined by (5.11) with $k=10$ in terms of the coefficients $C_{10}(d)$ given in Table 4.3.

A Siegel modular form sometimes also admits a product representation, and can be obtained as Borcherds lift or *multiplicative lift* of a weak Jacobi form of weight zero and index one. This procedure is in a sense an exponentiation of the additive lift and naturally results in the product representation of the Siegel form using the Fourier coefficients of a Jacobi form as the input. Several examples of Siegel forms that admit product representation are known but at present there is no general theory to determine under what conditions a given Siegel form admits a product representation.

For the Igusa cusp form Φ_{10}, a product representation does exist. It was obtained by Gritsenko and Nikulin [67, 68] as a multiplicative lift of the elliptic genus $\chi(\tau,z;K3) = 2\varphi_{0,1}(\tau,z)$ and is given by

$$\Phi_{10}(\Omega) \;=\; qyp \prod_{(r,s,t)>0} \big(1-q^s y^t p^r\big)^{2C_0(4rs-t^2)}, \tag{5.13}$$

in terms of C_0 given by (4.30, 4.24). Here the notation $(r,s,t)>0$ means that $r,\ s,\ t\ \in \mathbb{Z}$ with either $r>0$ or $r=0,\ s>0$, or $r=s=0,\ t<0$. In terms of the V_m, this can be rewritten in the form $\Phi_{10} = p\,\varphi_{10,1}\,\exp\!\big(-2\sum_{m=1}^{\infty}(\varphi_{0,1}|V_m)\,p^m\big)$.

5.2 Quantum black holes and Siegel modular forms

Siegel forms occur naturally in the context of counting of quarter-BPS dyons. The partition function for these dyons depends on three (complexified) chemical potentials (σ, τ, z), conjugate to the three T-duality invariant integers (m, n, ℓ) respectively and is given by

$$\mathrm{Z}(\Omega) \;=\; \frac{1}{\Phi_{10}(\Omega)} \;. \tag{5.14}$$

The product representation of the Igusa form is particularly useful for the physics application because it is closely related to the generating function for the elliptic genera of symmetric products of $K3$ introduced earlier. This is a consequence of the fact that the multiplicative lift of the Igusa form is obtained starting with the elliptic genus of $K3$ as the input. Comparing the product representation for the Igusa form (5.13) with (4.52), we get the relation:

$$\mathrm{Z}(\tau, z, \sigma) \;=\; \frac{1}{\Phi_{10}(\tau, z, \sigma)} \;=\; \frac{\widehat{Z}(\tau, z, \sigma)}{\varphi_{10,1}(\tau, z)} \;. \tag{5.15}$$

This relation to the elliptic genera of symmetric products of $K3$ has a deeper physical significance based on what is known as the 4d-5d lift [60]. The main idea is to use the fact that the geometry of the Kaluza-Klein monopole in the charge configuration (2.15) reduces to five-dimensional flat Minkowski spacetime in the limit when the radius of the circle $\tilde{S}^1$ goes to infinity. In this limit, the charge l corresponding to the momentum around this circle gets identified with the angular momentum l in five dimensions. Our charge configuration (2.15) then reduces essentially to the Strominger-Vafa black hole [118] with angular momentum [12] discussed in the previous subsection. Assuming that the dyon partition function does not depend on the moduli, we thus essentially relate $Z(\Omega)$ to $\hat{Z}(\Omega)$. The additional factor in (5.15) involving $\Phi_{10}(\sigma, \tau, z)$ comes from bound states of momentum n with the Kaluza-Klein monopole and from the center of mass motion of the Strominger-Vafa black hole in the Kaluza-Klein geometry [59, 42].

The Igusa cusp form has double zeros at $z = 0$ and its $Sp(2, \mathbb{Z})$ images. The partition function is therefore a *meromorphic* Siegel form (5.7) of weight -10 with double poles at these divisors. This fact is responsible for much of the interesting physics of wall-crossings in this context as we explain in the next section.

Using (5.12) or (5.13), one can compute the coefficients ψ_m in the Fourier-Jacobi expansion (1.1). The first few, multiplied by Δ and a suitable integer, are given by

$$\begin{aligned} \Delta\,\psi_{-1} &= A^{-1}\,, \\ \Delta\,\psi_0 &= 2\,A^{-1}B\,, \\ 4\,\Delta\,\psi_1 &= 9\,A^{-1}B^2 + 3E_4A\,, \\ 27\,\Delta\,\psi_2 &= 50\,A^{-1}B^3 + 48E_4AB + 10E_6A^2\,, \\ 384\,\Delta\,\psi_3 &= 475\,A^{-1}B^4 + 886E_4AB^2 + 360E_6A^2B + 199E_4^2A^3\,, \\ 72\,\Delta\,\psi_4 &= 51\,A^{-1}B^5 + 155E_4AB^3 + 93E_6A^2B^2 + 102E_4^2A^3B + 31E_4E_6A^4\,. \end{aligned} \tag{5.16}$$

(Here E_4, E_6, Δ are as in §3 and $A = \varphi_{-2,1}$, $B = \varphi_{0,1}$ as in §4.) The double zero of Φ_{10} at $z = 0$ is reflected by the denominator A in the $A^{-1}B^{m+1}$ terms in these formulas. Note that, by (5.15), the right-hand side of (5.16) multiplied by A and divided by a suitable denominator are equal to the elliptic genus of symmetric products of a $K3$ surface. For example,

$$\chi(\mathrm{Sym}^4K3;\tau,z) = \frac{1}{384}\big(475\,B^4 + 886E_4A^2B^2 + 360E_6A^3B + 199E_4^2A^4\,\big)\,. \tag{5.17}$$

6. Walls and contours

Given the partition function (5.14), one can extract the black hole degeneracies from the Fourier coefficients. The three quadratic T-duality invariants of a given dyonic state can be organized as a 2×2 symmetric matrix

$$\Lambda = \begin{pmatrix} N\cdot N & N\cdot M \\ M\cdot N & M\cdot M \end{pmatrix} = \begin{pmatrix} 2n & \ell \\ \ell & 2m \end{pmatrix}\,, \tag{6.1}$$

where the dot products are defined using the $O(22,6;\mathbb{Z})$ invariant metric L. The matrix Ω in (5.14) and (5.4) can be viewed as the matrix of complex chemical potentials conjugate to the charge matrix Λ. The charge matrix Λ is manifestly T-duality invariant. Under an S-duality transformation (2.4), it transforms as

$$\Lambda \rightarrow \gamma\Lambda\gamma^t \tag{6.2}$$

There is a natural embedding of this physical S-duality group $SL(2,\mathbb{Z})$ into $Sp(2,\mathbb{Z})$:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} (\gamma^t)^{-1} & \mathbf{0} \\ \mathbf{0} & \gamma \end{pmatrix} = \begin{pmatrix} d & -c & 0 & 0 \\ -b & a & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix} \in Sp(2,\mathbb{Z})\,. \tag{6.3}$$

The embedding is chosen so that $\Omega \to (\gamma^T)^{-1}\Omega\gamma^{-1}$ and $\mathrm{Tr}(\Omega \cdot \Lambda)$ in the Fourier integral is invariant. This choice of the embedding ensures that the physical degeneracies extracted from the Fourier integral are S-duality invariant if we appropriately transform the moduli at the same time as we explain below.

To specify the contours, it is useful to define the following moduli-dependent quantities. One can define the matrix of right-moving T-duality invariants

$$\Lambda_R = \begin{pmatrix} Q_R \cdot Q_R & Q_R \cdot P_R \\ P_R \cdot Q_R & P_R \cdot P_R \end{pmatrix} , \tag{6.4}$$

which depends both on the integral charge vectors N, M as well as the T-moduli μ. One can then define two matrices naturally associated to the S-moduli $S = S_1 + iS_2$ and the T-moduli μ respectively by

$$\mathcal{S} = \frac{1}{S_2}\begin{pmatrix} |S|^2 & S_1 \\ S_1 & 1 \end{pmatrix} \quad , \quad \mathcal{T} = \frac{\Lambda_R}{|\det(\Lambda_R)|^{\frac{1}{2}}} . \tag{6.5}$$

Both matrices are normalized to have unit determinant. In terms of them, we can construct the moduli-dependent 'central charge matrix'

$$\mathcal{Z} = |\det(\Lambda_R)|^{\frac{1}{4}} \left(\mathcal{S} + \mathcal{T}\right), \tag{6.6}$$

whose determinant equals the BPS mass

$$M_{Q,P} = |\det \mathcal{Z}| . \tag{6.7}$$

We define

$$\widetilde{\Omega} \equiv \begin{pmatrix} \sigma & -z \\ -z & \tau \end{pmatrix} . \tag{6.8}$$

This is related to Ω by an $SL(2,\mathbb{Z})$ transformation

$$\tilde{\Omega} = \widehat{S}\Omega\widehat{S}^{-1} \quad \text{where} \quad \widehat{S} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \tag{6.9}$$

so that, under a general S-duality transformation γ, we have the transformation $\tilde{\Omega} \to \gamma\widetilde{\Omega}\gamma^T$ as $\Omega \to (\gamma^T)^{-1}\Omega\gamma^{-1}$.

With these definitions, $\Lambda, \Lambda_R, \mathcal{Z}$ and $\tilde{\Omega}$ all transform as $X \to \gamma X \gamma^T$ under an S-duality transformation (2.4) and are invariant under T-duality transformations. The moduli-dependent Fourier contour can then be specified in a duality-invariant fashion by[26]

$$\mathcal{C} = \{\mathrm{Im}(\widetilde{\Omega}) = \varepsilon^{-1}\mathcal{Z}; \quad 0 \leq \mathrm{Re}(\tau),\, \mathrm{Re}(\sigma),\, \mathrm{Re}(z) < 1\}, \tag{6.10}$$

where $\varepsilon \to 0^+$. For a given set of charges, the contour depends on the moduli S, μ through the definition of the central charge vector (6.6). The degeneracies $d(n, \ell, m)|_{S,\mu}$ of states with the T-duality invariants (n, ℓ, m) at a given point (S, μ) in the moduli space are then given by[11]

$$d(n, \ell, m)|_{S,\mu} = \int_{\mathcal{C}} e^{-i\pi \mathrm{Tr}(\Omega \cdot \Lambda)} \; \mathrm{Z}(\Omega) \, d^3\Omega \,. \tag{6.11}$$

This contour prescription thus specifies how to extract the degeneracies from the partition function for a given set of charges and in any given region of the moduli space. In particular, it also completely summarizes all wall-crossings as one moves around in the moduli space for a fixed set of charges. Even though the indexed partition function has the same functional form throughout the moduli space, the spectrum is moduli dependent because of the moduli dependence of the contours of Fourier integration and the pole structure of the partition function. Since the degeneracies depend on the moduli *only* through the dependence of the contour $\mathcal{C}$, moving around in the moduli space corresponds to deforming the Fourier contour. This does not change the degeneracy except when one encounters a pole of the partition function. Crossing a pole corresponds to crossing a wall in the moduli space. The moduli space is thus divided up into domains separated by "walls of marginal stability." In each domain the degeneracy is constant but it jumps upon crossing a wall as one goes from one domain to the other. The jump in the degeneracy has a nice mathematical characterization. It is simply given by the residue at the pole that is crossed while deforming the Fourier contour in going from one domain to the other.

We now turn to the degeneracies of single-centered black holes. Given the T-duality invariants Λ, a single centered black hole solution is known to exist in all regions of the moduli space as long as $\det(\Lambda)$ is large and positive. The moduli fields can take any values $(\lambda_\infty, \mu_\infty)$ at asymptotic infinity far away from the black hole but the vary in the black hole geometry. Because of the attractor phenomenon [58, 116], the moduli adjust themselves so that near the horizon of the black hole of charge Λ they get attracted to the values $(\lambda_*(\Lambda), \mu_*(\Lambda))$ which are determined by the requirement that the central charge $\mathcal{Z}_*(\Lambda)$ evaluated using these moduli becomes proportional to Λ. These attractor values are independent of the asymptotic values and depend only on the charge of black hole. We call these moduli the attractor moduli. This enables us to define the *attractor contour* for a given charge Λ by fixing the asymptotic moduli to the attractor values corresponding to this charge. In this case

$$\mathcal{Z}(\lambda_\infty, \mu_\infty) \;=\; \mathcal{Z}(\lambda_*(\Lambda), \mu_*(\Lambda)) \sim \Lambda \tag{6.12}$$

and we have the attractor contour

$$\mathcal{C}_* \;=\; \{\mathrm{Im}(\widetilde{\Omega}) \;=\; \varepsilon^{-1}\Lambda; \quad 0 \leq \mathrm{Re}(\tau),\, \mathrm{Re}(\sigma),\, \mathrm{Re}(z) < 1\} \tag{6.13}$$

---

[11] The physical degeneracies have an additional multiplicative factor of $(-1)^{\ell+1}$ which we omit here for simplicity of notation in later chapters.

which depends only on the integral charges and not on the moduli. The significance of the attractor moduli in our context stems from the fact if the asymptotic moduli are tuned to these values for given (n, ℓ, m), then only single-centered black hole solution exists. The degeneracies $d^*(n, \ell, m)$ obtained using the attractor contour

$$d^*(n, \ell, m) = \int_{\mathcal{C}_*} e^{-i\pi \mathrm{Tr}(\Omega\cdot\Lambda)} \, \mathrm{Z}(\Omega) \, d^3\Omega \tag{6.14}$$

are therefore expected to be the degeneracies of the immortal single-centered black holes.

7. Mock modular forms

Mock modular forms are a relatively new class of modular objects, although individual examples had been known for some time. They were first isolated explicitly by S. Zwegers in his thesis [130] as the explanation of the "mock theta functions" introduced by Ramanujan in his famous last letter to Hardy. An expository account of this work can be found in [129].

Ramanujan's mock theta functions are examples of what we will call "pure" mock modular forms. By this we mean a holomorphic function in the upper half plane which transforms under modular transformations almost, but not quite, as a modular form. The non-modularity is of a very special nature and is governed by another holomorphic function called the *shadow* which is itself an ordinary modular form. We will describe this more precisely in §7.1 and give a number of examples. In §7.2, we introduce a notion of mock Jacobi forms (essentially, holomorphic functions of τ and z with theta expansions like that of usual Jacobi forms, but in which the coefficients $h_\ell(\tau)$ are mock modular forms) and show how all the examples given in the first subsection occur naturally as pieces of mock Jacobi forms. Finally in §7.3 we define a more general notion of mock modular forms in which the shadow is replaced by a sum of products of holomorphic and anti-holomorphic modular forms. This has the advantage that there are now many examples of strongly holomorphic mock modular forms, whereas pure mock modular forms almost always have a negative power of q at at least one cusp, and also that almost all of the examples occurring naturally in number theory, algebraic geometry, or physics are of this more general "mixed" type, with the pure mock modular forms arising as quotients of mixed mock modular forms by eta products.

7.1 Pure mock modular forms

We define a (weakly holomorphic) *pure mock modular form* of weight $k \in \frac{1}{2}\mathbb{Z}$ as the first member of a pair (h, g), where

1. h is a holomorphic function in $\mathbb{H}$ with at most exponential growth at all cusps,

2. the function $g(\tau)$, called the *shadow* of h, is a holomorphic[12] modular form of weight $2-k$, and
3. the sum $\widehat{h} := h + g^*$, called the *completion* of h, transforms like a holomorphic modular form of weight k, *i.e.* $\widehat{h}(\tau)/\theta(\tau)^{2k}$ is invariant under $\tau \to \gamma\tau$ for all $\tau \in \mathbb{H}$ and for all γ in some congruence subgroup of $SL(2,\mathbb{Z})$.

Here $g^*(\tau)$, called the *non-holomorphic Eichler integral*, is a solution of the differential equation

$$(4\pi\tau_2)^k \, \frac{\partial g^*(\tau)}{\partial\overline{\tau}} \;=\; -2\pi i \, \overline{g(\tau)} \, . \tag{7.1}$$

If g has the Fourier expansion $g(\tau) = \sum_{n\ge0} b_n \, q^n$, we fix the choice of g^* by setting

$$g^*(\tau) \;=\; \overline{b}_0 \, \frac{(4\pi\tau_2)^{-k+1}}{k-1} \;+\; \sum_{n>0} n^{k-1} \, \overline{b}_n \, \Gamma(1-k, 4\pi n\tau_2) \, q^{-n} \, , \tag{7.2}$$

where $\tau_2 = \mathrm{Im}(\tau)$ and $\Gamma(1-k,x) = \int_x^\infty t^{-k} \, e^{-t} \, dt$ denotes the incomplete gamma function, and where the first term must be replaced by $-\overline{b}_0 \, \log(4\pi\tau_2)$ if $k=1$. Note that the series in (7.2) converges despite the exponentially large factor q^{-n} because $\Gamma(1-k,x) = O(x^{-k}e^{-x})$. If we assume either that $k>1$ or that $b_0 = 0$, then we can define g^* alternatively by the integral

$$g^*(\tau) = \left(\frac{i}{2\pi}\right)^{k-1} \int_{-\overline{\tau}}^{\infty} (z+\tau)^{-k} \; \overline{g(-\overline{z})} \; dz \, . \tag{7.3}$$

(The integral is independent of the path chosen because the integrand is holomorphic in z.) Since h is holomorphic, (7.1) implies that the completion of h is related to its shadow by

$$(4\pi\tau_2)^k \, \frac{\partial\widehat{h}(\tau)}{\partial\overline{\tau}} \;=\; -2\pi i \, \overline{g(\tau)} \, . \tag{7.4}$$

We denote by $\mathbb{M}^!_{k|0}$ the space of weakly holomorphic pure mock modular forms of weight k and arbitrary level. (The "0" in the notation corresponds to the word "pure" and will be replaced by an arbitrary integer or half-integer in §7.3.) This space clearly contains the space $M^!_k$ of ordinary weakly holomorphic modular forms (the special case $g=0$, $h=\widehat{h}$) and we have an exact sequence

$$0 \longrightarrow M^!_k \longrightarrow \mathbb{M}^!_{k|0} \xrightarrow{\mathcal{S}} \overline{M_{2-k}} \;\; , \tag{7.5}$$

where the *shadow map* $\mathcal{S}$ sends h to $\overline{g}$.[13]

---

[12] One can also consider the case where the shadow is allowed to be a weakly holomorphic modular form, but we do not do this since none of our examples will be of this type.

[13] We will use the word "shadow" to denote either $g(\tau)$ or $\overline{g(\tau)}$, but the shadow *map*, which should be linear over $\mathbb{C}$, always sends h to $\overline{g}$, the complex conjugate of its holomorphic shadow. We will also often be sloppy and say that the shadow of a certain mock modular form "is" some modular form g when in fact it is merely proportional to g, since the constants occurring are arbitrary and are often messy.

In the special case when the shadow g is a unary theta series as in (3.10) or (3.11) (which can only hapen if k equals $3/2$ or $1/2$, respectively), the mock modular form h is called a *mock theta function.* All of Ramanujan's examples, and all of ours in this paper, are of this type. In these cases the incomplete gamma functions in (7.2) can be expressed in terms of the complementary error function $\operatorname{erfc}(x) = 2\pi^{-1/2}\int_x^\infty e^{-t^2}dt$:

$$\Gamma\bigl(-\frac{1}{2},\,x\bigr) \;=\; \frac{2}{\sqrt{x}}\,e^{-x} \;-\; 2\sqrt{\pi}\,\operatorname{erfc}\bigl(\sqrt{x}\bigr)\,, \qquad \Gamma\bigl(\frac{1}{2},\,x\bigr) \;=\; \sqrt{\pi}\,\operatorname{erfc}\bigl(\sqrt{x}\bigr)\,. \tag{7.6}$$

A very simple, though somewhat artificial, example of a mock modular form is the weight 2 Eisenstein series $E_2(\tau)$ mentioned in §3.1, which is also a quasimodular form. Here the shadow $g(\tau)$ is a constant and its non-holomorphic Eichler integral $g^*(\tau)$ simply a multiple of $1/\tau_2$. This example, however, is atypical, since most quasimodular forms (like $E_2(\tau)^2$) are not mock modular forms. We end this subsection by giving a number of less trivial examples. Many more will occur later.

Example 1. In Ramanujan's famous last letter to Hardy in 1920, he gives 17 examples of mock theta functions, though without giving any complete definition of this term. All of them have weight 1/2 and are given as q-hypergeometric series. A typical example (Ramanujan's second mock theta function of "order 7" — a notion that he also does not define) is

$$\mathcal{F}_{7,2}(\tau) = -q^{-25/168}\sum_{n=1}^{\infty}\frac{q^{n^2}}{(1-q^n)\cdots(1-q^{2n-1})} \;=\; -q^{143/168}\left(1+q+q^2+2q^3+\cdots\right)\,. \tag{7.7}$$

This is a mock theta function of weight 1/2 on $\Gamma_0(4)\cap\Gamma(7)$ with shadow the unary theta series

$$\sum_{n\equiv 2\,(\mathrm{mod}\ 7)} \chi_{12}(n)\,n\,q^{n^2/168}\,, \tag{7.8}$$

with $\chi_{12}(n)$ as in (3.13). The product $\eta(\tau)\mathcal{F}_{7,2}(\tau)$ is a strongly holomorphic mixed mock modular form of weight $(1,1/2)$, and by an identity of Hickerson [72] is equal to an indefinite theta series:

$$\eta(\tau)\,\mathcal{F}_{7,2}(\tau) \;=\; \sum_{r,\,s\,\in\,\mathbb{Z}+\frac{5}{14}} \frac{1}{2}\bigl(\operatorname{sgn}(r)+\operatorname{sgn}(s)\bigr)\,(-1)^{r-s}\,q^{(3r^2+8rs+3s^2)/2}\,. \tag{7.9}$$

Example 2. The second example is the generating function of the Hurwitz-Kronecker class numbers $H(N)$. These numbers are defined for $N>0$ as the number of $PSL(2,\mathbb{Z})$-equivalence classes of integral binary quadratic forms of discriminant $-N$, weighted by the reciprocal of the number of their automorphisms (if $-N$ is the discriminant of an imaginary quadratic field

K other than $\mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$, this is just the class number of K), and for other values of N by $H(0) = -1/12$ and $H(N) = 0$ for $N < 0$. It was shown in [126] that the function

$$\mathbf{H}(\tau) \ := \ \sum_{N=0}^{\infty} H(N)\, q^N \ = \ -\frac{1}{12} \ + \ \frac{1}{3}q^3 \ + \ \frac{1}{2}q^4 \ + \ q^7 \ + \ q^8 \ + \ q^{11} \ + \ \cdots \tag{7.10}$$

is a mock modular form of weight 3/2 on $\Gamma_0(4)$, with shadow the classical theta function $\theta(\tau) = \sum q^{n^2}$. Here $\mathbf{H}(\tau)$ is strongly holomorphic, but this is very exceptional. In fact, up to minor variations it is essentially the *only* known non-trivial example of a strongly holomorphic pure mock modular form, which is why we will introduce mixed mock modular forms below. As mentioned earlier, this mock modular form has appeared in the physics literature in the context of topologically twisted supersymmetric Yang-Mills theory on $\mathbb{CP}^2$ [120].

Example 3. This example is taken from [129]. Set

$$F_2^{(6)}(\tau) \ = \ -\sum_{r>s>0} \chi_{12}(r^2 - s^2)\, s\, q^{rs/6} \ = \ q + 2q^2 + q^3 + 2q^4 - q^5 + \cdots \tag{7.11}$$

with χ_{12} as in (3.13), and let $E_2(\tau)$ be the quasimodular Eisenstein series of weight 2 on $SL(2,\mathbb{Z})$ as defined in §3.1. Then the function

$$h^{(6)}(\tau) \ = \ \frac{12F_2^{(6)}(\tau) - E_2(\tau)}{\eta(\tau)} \ = \ q^{-1/24}\left(-1 + 35q + 130q^2 + 273q^3 + 595q^4 + \cdots\right) \tag{7.12}$$

is a weakly holomorphic mock modular form of weight 3/2 on $SL(2,\mathbb{Z})$ with shadow proportional to $\eta(\tau)$. More generally, if we define

$$F_k^{(6)}(\tau) \ = \ -\sum_{r>s>0} \chi_{12}(r^2 - s^2)\, s^{k-1} q^{rs/6} \qquad (k > 2,\ k \text{ even})\,, \tag{7.13}$$

then for all $\nu \geq 0$ we have

$$\frac{24^\nu}{\binom{2\nu}{\nu}}\, [h^{(6)}, \eta]_\nu \ = \ 12F_k^{(6)} \ - \ E_k \ + \ \text{cusp form of weight } k\ , \quad k = 2\nu + 2\ , \tag{7.14}$$

where $[h^{(6)}, \eta]_\nu$ denotes the νth Rankin-Cohen bracket of the mock modular form $h^{(6)}$ and the modular form η in weights $(3/2, 1/2)$. This statement, and the similar statements for other mock modular forms which come later, are proved by the method of holomorphic projection, which we do not explain here, and are intimately connected with the mock Jacobi forms introduced in the next subsection. That connection will also explain the superscript "6" in (7.11)–(7.14).

Example 4. Our last example is very similar to the preceding one. Set

$$F_2^{(2)}(\tau) \ = \ \sum_{\substack{r>s>0 \\ r-s \text{ odd}}} (-1)^r\, s\, q^{rs/2} \ = \ q + q^2 - q^3 + q^4 - q^5 + \cdots\,. \tag{7.15}$$

Then the function

$$h^{(2)}(\tau) \;=\; \frac{24F_2^{(2)}(\tau)-E_2(\tau)}{\eta(\tau)^3} \;=\; q^{-1/8}\left(-1+45q+231q^2+770q^3+2277q^4+\cdots\right) \tag{7.16}$$

is a weakly holomorphic mock modular form of weight $1/2$ on $SL(2,\mathbb{Z})$ with shadow proportional to $\eta(\tau)^3$ and, as in Example 3, if we set

$$F_k^{(2)}(\tau) \;=\; -\sum_{\substack{r>s>0\\ r-s\ \mathrm{odd}}}(-1)^r\, s^{k-1}\, q^{rs/2} \qquad (k>2,\ k\ \text{even})\,, \tag{7.17}$$

then for all $\nu \geq 0$ we have

$$\frac{8^\nu}{\binom{2\nu}{\nu}}\,[h^{(2)},\eta^3]_\nu \;=\; 24F_k^{(2)} \;-\; E_k \;+\; \text{cusp form of weight } k\ , \quad k=2\nu+2\ , \tag{7.18}$$

where $[h^{(2)},\eta^3]_\nu$ denotes the Rankin-Cohen bracket in weights $(1/2,3/2)$.

In fact, the mock modular form $h^{(2)}$, with the completely different definition

$$h^{(2)}(\tau) \;=\; \frac{\vartheta_2(\tau)^4-\vartheta_4(\tau)^4}{\eta(\tau)^3} \;-\; \frac{24}{\vartheta_3(\tau)}\sum_{n\in\mathbb{Z}}\frac{q^{n^2/2-1/8}}{1+q^{n-1/2}}\,, \tag{7.19}$$

arose from the works of Eguchi-Taormina [56] and Eguchi–Ooguri–Taormina–Yang [53] in the late 1980's in connection with characters of the $\mathcal{N}=4$ superconformal algebra in two dimensions and the elliptic genus of $K3$ surfaces, and appears explicitly in the work of Ooguri [99] (it is $q^{-1/8}(-1+F(\tau)/2)$ in Ooguri's notation) and Wendland [121]. It has recently aroused considerable interest because of the "Mathieu moonshine" discovery by Eguchi, Ooguri and Tachikawa (see [52, 24]) that the coefficients 45, 231, 770, … in (7.16) are the dimensions of certain representations of the Mathieu group M_{24}. The equality of the right-hand sides of (7.16) and (7.19), which is not a priori obvious, follows because both expressions define mock modular forms whose shadow is the same multiple of $\eta(\tau)^3$, and the first few Fourier coefficients agree.

7.2 Mock Jacobi forms

By a (pure) *mock Jacobi form*[14] (resp. *weak mock Jacobi form*) of weight k and index m we will mean a holomorphic function φ on $\mathbb{H}\times\mathbb{C}$ that satisfies the elliptic transformation property (4.2), and hence has a Fourier expansion as in (4.3) with the periodicity property (4.4) and a theta expansion as in (4.10), and that also satisfies the same cusp conditions (4.5) (resp. (4.7)) as in the classical case, but in which the modularity property with respect to the action of $SL(2,\mathbb{Z})$ on $\mathbb{H}\times\mathbb{Z}$ is weakened: the coefficients $h_\ell(\tau)$ in (4.10) are now mock modular forms rather than modular forms of weight $k-\frac{1}{2}$, and the modularity property of φ is that the *completed* function

---

[14] We mention that there are related constructions and definitions in the literature. The weak Maass-Jacobi forms of [14] include the completions of our mock Jacobi forms, but in general are non-holomorphic in z as well as in τ. See also [27] and [18]. The functions discussed in this paper will be holomorphic in z.

$$\widehat{\varphi}(\tau,z) \;=\; \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}} \widehat{h}_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z)\,, \tag{7.20}$$

rather than φ itself, transforms according to (4.1). If g_ℓ denotes the shadow of h_l, then we have

$$\widehat{\varphi}(\tau,z) \;= \varphi(\tau,z) \;+ \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}} g^*_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z)$$

with g^*_ℓ as in (7.2) and hence, by (7.1),

$$\psi(\tau,z) \;:=\; \tau_2^{k-1/2}\,\frac{\partial}{\partial\overline{\tau}}\widehat{\varphi}(\tau,z) \;\doteq\; \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}} \overline{g_\ell(\tau)}\,\vartheta_{m,\ell}(\tau,z)\,. \tag{7.21}$$

(Here $\doteq$ indicates an omitted constant.) The function $\psi(\tau,z)$ is holomorphic in z, satisfies the same elliptic transformation property (4.2) as φ does (because each $\vartheta_{m,\ell}$ satisfies this), satisfies the heat equation $\left(8\pi i m\,\frac{\partial}{\partial\tau}-\frac{\partial^2}{\partial z^2}\right)\psi=0$ (again, because each $\vartheta_{m,\ell}$ does), and, by virtue of the modular invariance property of $\widehat{\varphi}(\tau,z)$, also satisfies the transformation property

$$\psi\Big(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\Big) \;=\; |c\tau+d|\,(c\overline{\tau}+d)^{2-k}\,e^{\frac{2\pi i m c z^2}{c\tau+d}}\,\psi(\tau,z) \qquad \forall \begin{pmatrix} a & b\\ c & d\end{pmatrix} \in SL(2;\mathbb{Z}) \tag{7.22}$$

with respect to the action of the modular group. These properties say precisely that ψ is a *skew-holomorphic Jacobi form* of weight $3-k$ and index m in the sense of Skoruppa [112, 113], and the above discussion can be summarized by saying that we have an exact sequence

$$0 \longrightarrow J^{\text{weak}}_{k,m} \longrightarrow \mathbb{J}^{\text{weak}}_{k|0,m} \xrightarrow{\mathcal{S}} J^{\text{skew}}_{3-k,m} \tag{7.23}$$

(and similarly with the word "weak" omitted), where $\mathbb{J}^{\text{weak}}_{k|0,m}$ (resp. $\mathbb{J}_{k|0,m}$) denotes the space of weak (resp. strong) pure mock Jacobi forms and the "shadow map" $\mathcal{S}$ sends[15] φ to ψ.

It turns out that most of the classical examples of mock theta functions occur as the components of a vector-valued mock modular form which gives the coefficients in the theta series expansion of a mock Jacobi form. We illustrate this for the four examples introduced in the previous subsection.

Example 1. The function $\mathcal{F}_{7,2}(\tau)$ in the first example of §7.1 is actually one of three "order 7 mock theta functions" $\{\mathcal{F}_{7,j}\}_{j=1,2,3}$ defined by Ramanujan, each given by a q-hypergeometric formula like (7.7), each having a shadow $\Theta_{7,j}$ like in (7.8) (but now with the summation over

---

[15] This shadow map was introduced in the context of weak Maass-Jacobi forms in [14].

$n \equiv j$ rather than $n \equiv 2$ modulo 7), and each satisfying an indefinite theta series identity like (7.9). We extend $\{\mathcal{F}_{7,j}\}$ to all j by defining it to be an odd periodic function of j of period 7, so that the shadow of $\mathcal{F}_{7,j}$ equals $\Theta_{7,j}$ for all $j \in \mathbb{Z}$. Then the function

$$\mathcal{F}_{42}(\tau, z) \;=\; \sum_{\ell \pmod{84}} \chi_{12}(\ell)\, \mathcal{F}_{7,\ell}(\tau)\, \vartheta_{42,\ell}(\tau, z) \tag{7.24}$$

belongs to $\mathbb{J}^{\text{weak}}_{1,42}$. The Taylor coefficients ξ_ν as defined in equation (4.16) are proportional to $\sum_{j=1}^{3}\big[\mathcal{F}_{7,j}, \Theta_{7,j}\big]_\nu$ and have the property that their completions $\widehat{\xi}_\nu = \sum_{j=1}^{3}\big[\widehat{\mathcal{F}}_{7,j}, \Theta_{7,j}\big]_\nu$ transform like modular forms of weight $2\nu + 2$ on the full modular group $SL(2, \mathbb{Z})$.

Example 2. Set $\mathcal{H}_0(\tau) = \sum_{n=0}^{\infty} H(4n) q^n$ and $\mathcal{H}_1(\tau) = \sum_{n=1}^{\infty} H(4n-1) q^{n-\frac{1}{4}}$ so that $\mathcal{H}_0(\tau) + \mathcal{H}_1(t) = \mathbf{H}(\tau/4)$. Then the function

$$\mathcal{H}(\tau, z) \;=\; \mathcal{H}_0(\tau)\vartheta_{1,0}(\tau, z) \;+\; \mathcal{H}_1(\tau)\vartheta_{1,1}(\tau, z) \;=\; \sum_{\substack{n,\, r \in \mathbb{Z} \\ 4n - r^2 \geq 0}} H(4n - r^2)\, q^n\, y^r\,, \tag{7.25}$$

is a mock Jacobi form of weight 2 and index 1 with shadow $\overline{\vartheta_{1,0}(\tau,0)}\,\vartheta_{1,0}(\tau,z) + \overline{\vartheta_{1,1}(\tau,0)}\,\vartheta_{1,1}(\tau,z)$. The νth Taylor coefficient ξ_ν of $\mathcal{H}$ is given by

$$\frac{4^\nu}{\binom{2\nu}{\nu}} \sum_{j=0}^{1} [\vartheta_{1,j}, \mathcal{H}_j]_\nu \;=\; \delta_{k,2}\, E_k \;-\; F_k^{(1)} \;+\; (\text{cusp form of weight } k \text{ on } SL(2,\mathbb{Z}))\,, \tag{7.26}$$

where $k = 2\nu + 2$ and

$$F_k^{(1)}(\tau) \;:=\; \sum_{n>0} \Big(\sum_{d|n} \min\big(d, \frac{n}{d}\big)^{k-1} \Big) q^n \qquad (k \text{ even},\ k \geq 2)\,. \tag{7.27}$$

In fact the cusp form appearing in (7.26) is a very important one, namely (up to a factor -2) the sum of the normalized Hecke eigenforms in $S_k(SL(2,\mathbb{Z}))$, and equation (7.26) is equivalent to the famous formula of Eichler and Selberg expressing the traces of Hecke operators on $S_k(SL(2,\mathbb{Z}))$ in terms of class numbers of imaginary quadratic fields.

Example 3. Write the function $h^{(6)}$ defined in (7.12) as

$$h^{(6)}(\tau) \;=\; \sum_{\substack{\Delta \geq -1 \\ \Delta \equiv -1 \pmod{24}}} C^{(6)}(\Delta)\, q^{\Delta/24} \tag{7.28}$$

with the first few coefficients $C^{(6)}(\Delta)$ given by

Δ	-1	23	47	71	95	119	143	167	191	215	239
$C^{(6)}(\Delta)$	-1	35	130	273	595	1001	1885	2925	4886	7410	11466

.

Then the function

$$\mathcal{F}_6(\tau,z) \;=\; \sum_{\substack{n,\,r\in\mathbb{Z}\\ 24n-r^2\ge -1}} \chi_{12}(r)\,C^{(6)}(24n-r^2)\,q^n\,y^r \tag{7.29}$$

is a mock Jacobi form of index 6 (explaining the notation $h^{(6)}$). Note that, surprisingly, this is even *simpler* than the expansion of the index 1 mock Jacobi form just discussed, because its twelve Fourier coefficients h_ℓ are all multiples of just one of them, while the two Fourier coefficients h_ℓ of $\mathcal{H}(\tau,z)$ were not proportional. (This is related to the fact that the shadow $\eta(\tau)$ of $h^{(6)}(\tau)$ is a modular form on the full modular group, while the shadow $\theta(\tau)$ of $\mathcal{H}(\tau)$ is a modular form on a congruence subgroup.) Specifically, we have $h_\ell(\tau) = \chi_{12}(\ell)h^{(6)}(\tau)$ for all $\ell \in \mathbb{Z}/12\mathbb{Z}$, where χ_{12} is the character defined in (3.13), so that $\mathcal{F}_6$ has the factorization

$$\mathcal{F}_6(\tau,z) \;=\; h^{(6)}(\tau) \sum_{\ell\ (\mathrm{mod} 12)} \chi_{12}(\ell)\,\vartheta_{6,\ell}(\tau,z) \;=\; \eta(\tau)\,h^{(6)}(\tau)\,\frac{\vartheta_1(\tau,4z)}{\vartheta_1(\tau,2z)}\,. \tag{7.30}$$

Combining this with (4.17) and noting that $\vartheta^0_{6,1} - \vartheta^0_{6,5} = \vartheta^0_{6,11} - \vartheta^0_{6,7} = \eta$, we see that the functions described in (7.14) are proportional to the Taylor coefficients $\xi_\nu(\tau)$ of $\mathcal{F}_6(\tau,z)$.

Example 4. The fourth example is very similar. Write the mock modular form (7.16) as

$$h^{(2)}(\tau) \;=\; \sum_{\substack{\Delta\ge -1\\ \Delta\equiv -1\ (\mathrm{mod}\ 8)}} C^{(2)}(\Delta)\,q^{\Delta/8} \tag{7.31}$$

with initial coefficients given by

Δ	-1	7	15	23	31	39	47	55	63	71
$C^{(2)}(\Delta)$	-1	45	231	770	2277	5796	13915	30843	65550	132825

.

Then the function

$$\begin{aligned}\mathcal{F}_2(\tau,z) \;&=\; \sum_{\substack{n,\,r\in\mathbb{Z}\\ 8n-r^2\ge -1}} \chi_4(r)\,C^{(2)}(8n-r^2)\,q^n\,y^r \;=\; h^{(2)}(\tau)\,\big(\vartheta_{2,1}(\tau,z)-\vartheta_{2,3}(\tau,z)\big)\\ &=\; h^{(2)}(\tau)\,\vartheta_1(\tau,2z) = \eta(\tau)^3\,h^{(2)}(\tau)\,C(\tau,z)\end{aligned} \tag{7.32}$$

(where $\chi_4(r) = \pm1$ for $r \equiv \pm1\,(\mathrm{mod}\ 4)$ and $\chi_4(r) = 0$ for r even) is a mock Jacobi form of index 2. (Here C is the Jacobi form of weight -1 and index 2 introduced in §4.3.) As in Example 3, the functions given in (7.18) are proportional to the Taylor coefficients ξ_ν of $\mathcal{F}_2$, because $\vartheta^1_{2,1} = -\vartheta^1_{2,3} = \eta^3$, where $\vartheta^1_{2,\ell}$ is defined by (4.18).

7.3 Mock modular forms: the general case

If we replace the condition "exponential growth" in 1. of 7.1 by "polynomial growth," we get the class of *strongly holomorphic mock modular forms*, which we can denote $\mathbb{M}_{k|0}$, and an exact sequence $0 \to M_k \to \mathbb{M}_{k|0} \to \overline{M_{2-k}}$. This is not very useful, however, because there are almost no examples of "pure" mock modular forms that are strongly holomorphic, essentially the only ones being the function $\mathcal{H}$ of Example 2 of §7.1 and its variants. It becomes useful if we generalize to *mixed mock modular forms* of mixed weight $k|\ell$. (Here k and ℓ can be integers or half-integers.) These are holomorphic functions $h(\tau)$, with polynomial growth at the cusps, that have completions $\widehat{h}$ of the form $\widehat{h} = h + \sum_j f_j\, g_j^*$ with $f_j \in M_\ell$, $g_j \in M_{2-k+\ell}$ that transform like modular forms of weight k. The space $\mathbb{M}_{k|\ell}$ of such forms thus fits into an exact sequence

$$0 \longrightarrow M_k \longrightarrow \mathbb{M}_{k|\ell} \xrightarrow{\mathcal{S}} M_\ell \otimes \overline{M_{2-k+\ell}}\,, \tag{7.33}$$

where the shadow map $\mathcal{S}$ now sends h to $\sum_j f_j\, \overline{g}_j$. If $\ell = 0$ this reduces to the previous definition, since $M_0 = \mathbb{C}$, but with the more general notion of mock modular forms there are now plenty of strongly holomorphic examples, and, just as for ordinary modular forms, they have much nicer properties (notably, polynomial growth of their Fourier coefficients) than the weakly holomorphic ones. If the shadow of a mixed mock modular form $h \in \mathbb{M}_{k|\ell}$ happens to contain only one term $f(\tau)\overline{g(\tau)}$, and if $f(\tau)$ has no zeros in the upper half-plane, then $f^{-1}\,h$ is a weakly holomorphic pure mock modular form of weight $k-\ell$ (and in fact, all weakly holomorphic pure mock modular forms arise in this way). For functions in the larger space $\mathbb{M}^!_{k|\ell}$ of weakly holomorphic mixed mock modular forms, defined in the obvious way, this always happens, so $\mathbb{M}^!_{k|\ell}$ simply coincides with $M_\ell \otimes \mathbb{M}^!_{k-\ell}$ and does not give anything new, but the smaller space $\mathbb{M}_{k|\ell}$ does not have such a decomposition and is more interesting. Moreover, there are many examples. In fact, apart from Ramanujan's original examples of mock theta functions, which were defined as q-hypergeometric series and were weakly holomorphic pure mock modular forms of weight 1/2, most of the examples of mock modular forms occurring "in nature," such as Appell-Lerch sums or indefinite theta series, are strongly holomorphic mixed modular forms.

We can also define "even more mixed" mock modular forms by replacing $\mathbb{M}_{k|\ell}$ by the space $\mathbb{M}_k = \sum_\ell \mathbb{M}_{k|\ell}$, i.e., by allowing functions whose shadow is a finite sum of products $f_j(\tau)\overline{g_j(\tau)}$ with the f_j of varying weights ℓ_j and g_j of weight $2-k+\ell_j$. This space fits into an exact sequence

$$0 \longrightarrow M_k \longrightarrow \mathbb{M}_k \xrightarrow{\mathcal{S}} \bigoplus_\ell M_\ell \otimes \overline{M_{2-k+\ell}}\;. \tag{7.34}$$

In fact, the functions in this space are the most natural objects, and from now on we will use "mock modular forms" to refer to this more general class, including the adjectives "mixed" and "pure" only when needed for emphasis. To justify this claim, we consider the space $\mathfrak{M}_k$ of all *real-analytic modular forms of weight* k, i.e., real-analytic functions that transform like modular forms of weight k and have polynomial growth at all cusps. More generally, one can consider the space $\mathfrak{M}_{k,\ell}$ of real-analytic functions of polynomial growth that transform according to $F(\gamma\tau) = (c\tau+d)^k(c\overline{\tau}+d)^\ell F(\tau)$ for all matrices $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ belonging to some subgroup of finite index of $SL(2,\mathbb{Z})$. Since the function $\tau_2 = \mathrm{Im}(\tau)$ transforms by $(\gamma\tau)_2 = \tau_2/|c\tau+d|^2$, the spaces $\mathfrak{M}_{k,\ell}$ and $\tau_2^r\,\mathfrak{M}_{k+r,\ell+r}$ coincide for all r. (Here, k, ℓ and r can each be integral or half-integral.) Obviously the space $\mathfrak{M}_{k,\ell}$ contains the space $M_k \otimes \overline{M_\ell}$, so by the comment just made, it also contains the direct sum $\bigoplus_{r\in\mathbb{Z}} \tau_2^r\, M_{k+r} \otimes \overline{M_{\ell+r}}$. Elements of this subspace will be called *decomposable*. Now, if a function $F \in \mathfrak{M}_k = \mathfrak{M}_{k,0}$ has the property that its $\overline{\tau}$-derivative (which automatically belongs to $\mathfrak{M}_{k,2}$) is decomposable, say as $\sum_j {\tau_2}^{r_j}\, f_j\, \overline{g_j}$ (where the weights of f_j and g_j are necessarily equal to $k+r_j$ and $2+r_j$, respectively), then F is the sum of a holomorphic function h and a multiple of $\sum_j f_j g_j^*$ and the function h is a (mixed) mock modular form in the sense just explained. Conversely, a holomorphic function h is a (mixed) mock modular form of weight k if and only if there exist numbers r_j and modular forms $f_j \in M_{k+r_j}$ and $g_j \in M_{2+r_j}$ such that the sum $\widehat{h} := h + \sum_j f_j g_j^*$ belongs to $\mathfrak{M}_k$. We can thus summarize the entire discussion by the following very simple definition:

> *A mock modular form is the holomorphic part of a real-analytic modular form whose $\overline{\tau}$-derivative is decomposable.*

As examples of this more general class we mention the functions $F_2^{(6)}$ and $F_2^{(2)}$ defined in equations (7.11) and (7.15), which are both (strong) mixed mock modular forms of total weight 2 and trivial character on the full modular group $SL(2,\mathbb{Z})$, but with different mixed weights: $12F_2^{(6)}$ is the sum of the functions $E_2(\tau)$ and $\eta(\tau)\,h^{(6)}(\tau)$ of weights $2|0$ and $\frac{3}{2}|\frac{1}{2}$, while $24F_2^{(2)}$ is the sum of the functions $E_2(\tau)$ and $\eta(\tau)^3 h^{(2)}(\tau)$ of weights $2|0$ and $\frac{1}{2}|\frac{3}{2}$.

We remark that in the pure case we have $\tau_2^k \cdot \partial_{\overline{\tau}} F \in \overline{M_{2-k}} = \mathrm{Ker}\big(\partial_\tau : \mathfrak{M}_{0,2-k} \to \mathfrak{M}_{2,2-k}\big)$, so that the subspace $\mathbb{M}_{k,0}$ of $\mathbb{M}_k$ can be characterized by a *single* second-order differential equation $\Delta_k \widehat{h} = 0$, where $\Delta_k = \tau_2^{2-k}\partial_\tau \tau_2^k \partial_{\overline{\tau}}$ is the Laplacian in weight k. For this reason, (weak) pure mock modular forms can be, and often are, viewed as the holomorphic parts of (weak) harmonic Maass forms, i.e., of functions in $\mathfrak{M}_k^!$ annihilated by Δ_k. But in the general case, there is no connection between mock modular forms or their completions and any second-order operators. Only the Cauchy-Riemann operator $\partial/\partial\overline{\tau}$ is needed, and the new definition of mock modular forms as given above is both more flexible and more readily generalizable than the more familiar definition of pure mock modular forms in terms of harmonic Maass forms.

Finally, we can define mixed mock Jacobi forms in exactly the same way as we defined pure ones (i.e., they have theta expansions in which the coefficients are mixed mock modular forms, with completions that transform like ordinary Jacobi forms). We will see in the next section that such objects arise very naturally as the "finite parts" of meromorphic Jacobi forms.

7.4 Superconformal characters and mock modular forms

As already mentioned, mock modular forms first made their appearance in physics some ten years before the work of Zwegers, in connection with the characters of superconformal algebas [56, 53]. The representations of the $\mathcal{N}=4$ superconformal algeba with central charge $c=6k$ are labeled by two quantum numbers, the energy h and spin l of the lowest weight state $|\Omega\rangle$, which satisfy $h>k/4$, $l\in\{1/2,1,\ldots,k/2\}$ for the non-BPS (massive) case and $h=k/4$, $l\in\{0,1/2,\ldots,k/2\}$ for the BPS (massless) case. The corresponding character, defined by

$$\mathrm{ch}^{\tilde{R}}_{k,h,l}(\tau,z) \;=\; \mathrm{Tr}\big(q^{L_0-c/24}\,e^{4\pi i z J_0}\big) \qquad \big(\,L_0|\Omega\rangle=h|\Omega\rangle, \quad J_0|\Omega\rangle=l|\Omega\rangle\big)\,,$$

turns out to be, up to a power of q, a Jacobi form in the non-BPS case but a "mock" object in the BPS case. In particular, for $k=1$ the non-BPS characters are given by

$$\mathrm{ch}^{\tilde{R}}_{1,h,\frac12}(\tau,z) \;=\; -\,q^{h-\frac38}\,\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^3} \;=\; -\,q^{h-\frac38}\,\eta(\tau)^3\,\varphi_{-2,1}(\tau,z) \qquad \big(h>\frac14\big) \tag{7.35}$$

and the BPS characters by

$$\mathrm{ch}^{\tilde{R}}_{1,\frac14,0}(\tau,z) \;=\; -\,\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^3}\,\mu(\tau,z)\,, \qquad \mathrm{ch}^{\tilde{R}}_{1,\frac14,\frac12}(z,\tau)+2\mathrm{ch}^{\tilde{R}}_{1,\frac14,0}(\tau,z) \;=\; -\,q^{-\frac18}\,\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^3}\,, \tag{7.36}$$

where $\mu(\tau,z)$ is the function defined by

$$\mu(\tau,z) \;=\; \frac{e^{\pi i z}}{\vartheta_1(z,\tau)}\sum_{n\in\mathbb{Z}}\frac{(-1)^n q^{\frac{n^2+n}{2}}e^{2\pi i n z}}{1\,-\,q^n\,e^{2\pi i z}} \qquad \big(q=e^{2\pi i\tau},\; y=e^{2\pi i z}\big)\,, \tag{7.37}$$

an "Appell-Lerch sum" of a type that also played a central role in Zwegers's work[16] and that we will discuss in the next section. The "mock" aspect is the fact that $\mu(\tau,z)$ must be completed by an appropriate non-holomorphic correction term in order to transform like a Jacobi form.

The mock modular form $h^{(2)}$ arises in connection with the decomposition of the elliptic genus of a $K3$ surface as a linear combination of these characters. This elliptic genus is, as discussed in 4.5, equal to the weak Jacobi form $2B=2\varphi_{0,1}$. It can be represented as a specialization of the partition function (= generating function of states) of superstrings compactified on a $K3$ surface [53] and as such can be decomposed into pieces corresponding to the representations of the $N=4$ superconformal algebra with central charge $6k=6$. The decomposition has the form [99, 121, 55]

$$2\varphi_{0,1}(\tau,z) \;=\; 20\,\mathrm{ch}^{\tilde{R}}_{1,\frac14,0}(z,\tau)\;+\;A_0\,\mathrm{ch}^{\tilde{R}}_{1,\frac14,\frac12}(z,\tau)\;+\;\sum_{n=1}^{\infty}A_n\,\mathrm{ch}^{\tilde{R}}_{1,n+\frac14,\frac12}(z,\tau)\,, \tag{7.38}$$

---

[16] More precisely, the function $\mu(\tau,z)$ is the specialization to $u=v=z$ of the 2-variable function $\mu(u,v;\tau)$ considered by Zwegers.

where $A_0 = -2$, $A_1 = 90$, $A_2 = 462$, $\ldots$ are the Fourier coefficients $2C^{(2)}(8n-1)$ of the mock modular form $2h^{(2)}(\tau)$. Inserting (7.35) and (7.36), we can rewrite this as

$$\eta(\tau)^{-3}\,\frac{\varphi_{0,1}(\tau,z)}{\varphi_{-2,1}(\tau,z)} \;=\; -\,12\,\mu(\tau,z) \,-\, h^{(2)}(\tau)\,. \tag{7.39}$$

In the next section we will see how to interpret this equation as an example of a certain canonical decomposition of meromorphic Jacobi forms (8.52).

8. From meromorphic Jacobi forms to mock modular forms

In this section we consider Jacobi forms $\varphi(\tau,z)$ that are meromorphic with respect to the variable z. It was discovered by Zwegers [130] that such forms, assuming that their poles occur only at points $z \in \mathbb{Q}\tau + \mathbb{Q}$ (i.e., at torsion points on the elliptic curve $\mathbb{C}/\mathbb{Z}\tau + \mathbb{Z}$), have a modified theta expansion related to mock modular forms. Our treatment is based on his, but the presentation is quite different and the results go further in one key respect. We show that φ decomposes canonically into two pieces, one constructed directly from its poles and consisting of a finite linear combination of Appell-Lerch sums with modular forms as coefficients, and the other being a mock Jacobi form in the sense introduced in the preceding section. Each piece separately transforms like a Jacobi form with respect to elliptic transformations. Neither piece separately transforms like a Jacobi form with respect to modular transformations, but each can be completed by the addition of an explicit and elementary non-holomorphic correction term so that it does transform correctly with respect to the modular group.

In §8.1 we explain how to modify the Fourier coefficients h_ℓ defined in (4.8) when φ has poles, and use these to define a "finite part" of φ by the theta decomposition (4.10). In §8.2 we define (in the case when φ has simple poles only) a "polar part" of φ as a finite linear combination of standard Appell-Lerch sums times modular forms arising as the residues of φ at its poles, and show that φ decomposes as the sum of its finite part and its polar part. Subsection 8.3 gives the proof that the finite part of φ is a mock Jacobi form and a description of the non-holomorphic correction term needed to make it transform like a Jacobi form. This subsection also contains a summary in tabular form of the various functions that have been introduced and the relations between them. In §8.4 we describe the modifications needed in the case of double poles (the case actually needed in this paper) and in §8.5 we present a few examples to illustrate the theory. Among the "mock" parts of these are two of the most interesting mock Jacobi forms from §7 (the one related to class numbers of imaginary quadratic fields and the one conjecturally related to representations of the Mathieu group M_{24}). Many other examples will be given in §9.

Throughout the following sections, we use the convenient notation $\mathbf{e}(x) := e^{2\pi i x}$.

8.1 The Fourier coefficients of a meromorphic Jacobi form

As indicated above, the main problem we face is to find an analogue of the theta decomposition (4.10) of holomorphic Jacobi forms in the meromorphic case. We will approach this problem from two sides: computing the Fourier coefficients of $\varphi(\tau, z)$ with respect to z, and computing the contribution from the poles. In this subsection we treat the first of these questions.

Consider a meromorphic Jacobi form $\varphi(\tau, z)$ of weight k and index m. We assume that $\varphi(\tau, z)$ for each $\tau \in \mathbb{H}$ is a meromorphic function of z which has poles only at points $z = \alpha\tau + \beta$ wit α and β rational. In the case when φ was holomorphic, we could write its Fourier expansion in the form (4.8). By Cauchy's theorem, the coefficient $h_\ell(\tau)$ in that expansion could also be given by the integral formula

$$h_\ell^{(z_0)}(\tau) \;=\; q^{-\ell^2/4m} \int_{z_0}^{z_0+1} \varphi(\tau, z)\, \mathbf{e}(-\ell z)\, dz\,, \tag{8.1}$$

where z_0 is an arbitrary point of $\mathbb{C}$. From the holomorphy and transformation properties of φ it follows that the value of this integral is independent of the choice of z_0 and of the path of integration and depends only on ℓ modulo $2m$ (implying that we have the theta expansion (4.10)) and that each h_ℓ is a modular form of weight $k - \frac{1}{2}$. Here each of these properties fails: the integral (8.1) is not independent of the path of integration (it jumps when the path crosses a pole); it is not independent of the choice of the initial point z_0; it is not periodic in ℓ (changing ℓ by $2m$ corresponds to changing z_0 by τ); it is not modular; and of course the expansion (4.10) cannot possibly hold since the right-hand-side has no poles in z.

To take care of the first of these problems, we specify the path of integration in (8.1) as the horizontal line from z_0 to $z_0 + 1$. If there are poles of $\varphi(\tau, z)$ along this line, this does not make sense; in that case, we define the value of the integral as the average of the integral over a path deformed to pass just above the poles and the integral over a path just below them. (We do not allow the initial point z_0 to be a pole of φ, so this makes sense.) To take care of the second and third difficulties, the dependence on z_0 and the non-periodicity in ℓ, we play one of these problems off against the other. From the elliptic transformation property (4.2) we find that

$$\begin{aligned} h_\ell^{(z_0+\tau)}(\tau) \;&=\; q^{-\ell^2/4m} \int_{z_0}^{z_0+1} \varphi(\tau,\, z+\tau)\, \mathbf{e}(-\ell\,(z+\tau))\, dz \\ &=\; q^{-(\ell+2m)^2/4m} \int_{z_0}^{z_0+1} \varphi(\tau,\, z)\, \mathbf{e}(-(\ell+2m)\, z)\, dz \;=\; h_{\ell+2m}^{(z_0)}(\tau)\,, \end{aligned}$$

i.e., changing z_0 by τ corresponds to changing ℓ by $2m$, as already mentioned. It follows that if we choose z_0 to be $-\ell\tau/2m$ (or $-\ell\tau/2m + B$ for any $B \in \mathbb{R}$, since it is clear that the value of the integral (8.1) depends only on the height of the path of integration and not on the initial point on this line), then the quantity

$$h_\ell(\tau) \;:=\; h_\ell^{(-\ell\tau/2m)}(\tau)\,, \tag{8.2}$$

which we will call the ℓth **canonical Fourier coefficient** of φ, depends only on the value of $\ell \pmod{2m}$. This in turn implies that the sum

$$\varphi^F(\tau,z) \;:=\; \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}} h_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z)\,, \tag{8.3}$$

which we will call the **finite** (or **Fourier**) **part** of φ, is well defined. If φ is holomorphic, then of course $\varphi^F = \varphi$, by virtue of (4.10). Hence if we define the **polar part** φ^P of φ to be $\varphi - \varphi^F$, then φ^P only depends on the principal part of φ at its singularities. We will give explicit formulas in §8.2 and §8.4 in the cases when all poles are single or double.

Note that the definiton of $h_\ell(\tau)$ can also be written

$$h_\ell(\tau) \;=\; q^{\ell^2/4m} \int_{\mathbb{R}/\mathbb{Z}} \varphi(\tau, z - \ell\tau/2m)\,\mathbf{e}(-\ell z)\,dz\,, \tag{8.4}$$

with the same convention as above if $\varphi(\tau, z - \ell\tau/2m)$ has poles on the real line.

Finally, we remark that none of these definitions have anything to do with the modular transformation property (4.1), and would also work if our initial function $\varphi(\tau,z)$ satisfied only the elliptic transformation property (4.2). (In that case, however, the functions h_ℓ in general need not be modular, and in fact will be mock modular when $\varphi(\tau,z)$ is a meromorphic Jacobi form.) What we have said so far can be summarized by a commutative diagram

$$\begin{tikzcd} \mathcal{E}^{\mathrm{Mer}}_{\pm,m} \arrow[r, "F"] \arrow[d, phantom, "\cup"] & \mathcal{E}_{\pm,m} \arrow[d, phantom, "\cup"] \\ J^{\mathrm{Mer}}_{k,m} \arrow[r, "F"] & \mathbb{J}_{k,m} \end{tikzcd}\;, \tag{8.5}$$

where $J^{\mathrm{Mer}}_{k,m}$ (resp. $\mathcal{E}^{\mathrm{Mer}}_{\pm,m}$) denote the subspace of meromorphic Jacobi (resp. elliptic) forms having poles only at torsion points $z \in \mathbb{Q}\tau + \mathbb{Q}$, and the map F sends $\varphi \mapsto \varphi^{\mathrm{F}}$.

8.2 The polar part of φ (case of simple poles)

We now consider the contribution from the poles. To present the results we first need to fix notations for the positions and residues of the poles of our meromorphic function φ. We assume for now that the poles are all simple.

By assumption, $\varphi(\tau,z)$ has poles only at points of the form $z = z_s = \alpha\tau + \beta$ for $s = (\alpha,\beta)$ belonging to some subset S of $\mathbb{Q}^2$. The double periodicity property (4.2) implies that S is invariant under translation by $\mathbb{Z}^2$, and of course $S/\mathbb{Z}^2$ must be finite. The modular transformation property (4.1) of φ implies that S is $SL(2,\mathbb{Z})$-invariant. For each $s = (\alpha,\beta) \in S$, we set

$$D_s(\tau) \;=\; 2\pi i\,\mathbf{e}(m\alpha z_s)\,\mathrm{Res}_{z=z_s}\big(\varphi(\tau,z)\big) \qquad (s = (\alpha,\beta) \in S,\; z_s = \alpha\tau + \beta)\,. \tag{8.6}$$

The functions $D_s(\tau)$ are holomorphic modular forms of weight $k-1$ and some level, and only finitely many of them are distinct. More precisely, they satisfy

- $$D_{(\alpha+\lambda,\beta+\mu)}(\tau) \;=\; \mathbf{e}(m(\mu\alpha-\lambda\beta))\, D_{(\alpha,\beta)}(\tau) \quad \text{for} \quad (\lambda,\mu) \,\in\, \mathbb{Z}^2\,, \tag{8.7}$$
- $$D_s\Big(\frac{a\tau+b}{c\tau+d}\Big) \;=\; (c\tau+d)^{k-1}\, D_{s\gamma}(\tau) \quad \text{for} \quad \gamma = \big(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\big) \,\in\, SL(2,\mathbb{Z})\,, \tag{8.8}$$

as one sees from the transformation properties of φ. (The calculation for equation (8.8) is given in §8.4.) It is precisely to obtain the simple transformation equation (8.8) that we included the factor $\mathbf{e}(m\alpha z_s)$ in (8.6). Since we are assuming for the time being that there are no higher-order poles, all of the information about the non-holomorphy of φ is contained in these functions.

The strategy is to define a "polar part" of φ by taking the poles z_s in some fundamental parallelogram for the action of the lattice $\mathbb{Z}\tau+\mathbb{Z}$ on $\mathbb{C}$ (i.e., for $s=(\alpha,\beta)$ in the intersection of S with some square box $[A,A+1)\times[B,B+1)$) and then averaging the residues at these poles over all translations by the lattice. But we must be careful to do this in just the right way to get the desired invariance properties. For each $m\in\mathbb{N}$ we introduce the *averaging operator*

$$\mathrm{Av}^{(m)}\big[F(y)\big] \;:=\; \sum_{\lambda} q^{m\lambda^2} y^{2m\lambda} F(q^\lambda y) \tag{8.9}$$

which sends any function of y (= $\mathbb{Z}$-invariant function of z) of polynomial growth in y to a function of z transforming like an index m Jacobi form under translations by the full lattice $\mathbb{Z}\tau+\mathbb{Z}$. For example, we have

$$q^{\ell^2/4m}\,\mathrm{Av}^{(m)}\big[y^\ell\big] \;=\; \sum_{\lambda\,\in\,\mathbb{Z}} q^{(\ell+2m\lambda)^2/4m}\, y^{\ell+2m\lambda} \;=\; \vartheta_{m,\ell}(\tau,z) \tag{8.10}$$

for any $\ell\in\mathbb{Z}$. If $F(y)$ itself is given as the average

$$F(y) \;=\; \mathrm{Av}_{\mathbb{Z}}\big[f(z)\big] \;:=\; \sum_{\mu\in\mathbb{Z}} f(z+\mu) \qquad (z\in\mathbb{C},\;\; y=\mathbf{e}(z)) \tag{8.11}$$

of a function $f(z)$ in $\mathbb{C}$ (of sufficiently rapid decay at infinity), then we have

$$\mathrm{Av}^{(m)}\big[F(y)\big] \;=\; \mathrm{Av}^{(m)}_{\mathbb{Z}\tau+\mathbb{Z}}\big[f(z)\big] \;:=\; \sum_{\lambda,\,\mu\in\mathbb{Z}} e^{2\pi i m(\lambda^2\tau+2\lambda z)}\, f(z+\lambda\tau+\mu)\,. \tag{8.12}$$

We want to apply the averaging operator (8.9) to the product of the function $D_s(\tau)$ with a standard rational function of y having a simple pole of residue 1 at $y=y_s=\mathbf{e}(z_s)$, but the choice of this rational function is not obvious. The right choice turns out to be $\mathcal{R}_{-2m\alpha}(y)$, where $\mathcal{R}_c(y)$ for $c\in\mathbb{R}$ is defined by

$$\mathcal{R}_c(y) \;=\; \begin{cases} \dfrac{1}{2}\, y^c\, \dfrac{y+1}{y-1} & \text{if } c\in\mathbb{Z}\,, \\[2ex] y^{\lceil c\rceil}\, \dfrac{1}{y-1} & \text{if } c\in\mathbb{R}\smallsetminus\mathbb{Z}\,. \end{cases} \tag{8.13}$$

(Here $\lceil c \rceil$ denotes the "ceiling" of c, i.e., the smallest integer $\geq c$. The right-hand side can also be written more uniformly as $\dfrac{1}{2}\,\dfrac{y^{\lfloor c \rfloor + 1} + y^{\lceil c \rceil}}{y-1}$, where $\lfloor c \rfloor = -\lceil -c \rceil$ denotes the "floor" of c, i.e., the largest integer $\leq c$.) This function looks artificial, but is in fact quite natural. First of all, by expanding the right-hand side of (8.13) in a geometric series we find

$$\mathcal{R}_c(y) \;=\; \begin{cases} -\sum^*_{\ell \geq c} y^\ell & \text{if } |y| < 1, \\ \;\;\sum^*_{\ell \leq c} y^\ell & \text{if } |y| > 1, \end{cases} \tag{8.14}$$

where the asterisk on the summation sign means that the term $\ell = c$ is to be counted with multiplicity $1/2$ when it occurs (which happens only for $c \in \mathbb{Z}$, explaining the case distinction in (8.13)). This formula, which can be seen as the prototypical example of wall-crossing, can also be written in terms of z as a Fourier expansion (convergent for all $z \in \mathbb{C} \smallsetminus \mathbb{R}$)

$$\mathcal{R}_c(\mathbf{e}(z)) \;=\; -\sum_{\ell \in \mathbb{Z}} \frac{\operatorname{sgn}(\ell - c) \,+\, \operatorname{sgn}(z_2)}{2}\, \mathbf{e}(\ell z) \qquad (y = \mathbf{e}(z),\;\; z_2 = \operatorname{Im}(z) \neq 0)\,, \tag{8.15}$$

without any case distinction. Secondly, $R_c(y)$ can be expressed in a natural way as an average:

Proposition 8.1. *For $c \in \mathbb{R}$ and $z \in \mathbb{C} \smallsetminus \mathbb{Z}$ we have*

$$\mathcal{R}_c(\mathbf{e}(z)) \;=\; Av_{\mathbb{Z}}\Big[\frac{\mathbf{e}(cz)}{2\pi i z}\Big]\,. \tag{8.16}$$

Proof: If $c \in \mathbb{Z}$, then

$$\operatorname{Av}_{\mathbb{Z}}\Big[\frac{\mathbf{e}(cz)}{2\pi i z}\Big] \;=\; \frac{y^c}{2\pi i} \sum_{n \in \mathbb{Z}} \frac{1}{z-n} \;=\; \frac{y^c}{2\pi i}\,\frac{\pi}{\tan \pi z} \;=\; \frac{y^c}{2}\,\frac{y+1}{y-1}$$

by a standard formula of Euler. If $c \notin \mathbb{Z}$ then the Poisson summation formula and (8.15) give

$$\begin{aligned} \operatorname{Av}_{\mathbb{Z}}\Big[\frac{\mathbf{e}(cz)}{2\pi i z}\Big] \;&=\; \sum_{n \in \mathbb{Z}} \frac{\mathbf{e}(c(z+n))}{2\pi i (z+n)} \;=\; \sum_{\ell \in \mathbb{Z}} \bigg(\int_{iz_2 - \infty}^{iz_2 + \infty} \frac{\mathbf{e}(c(z+u))}{2\pi i(z+u)}\, \mathbf{e}(-\ell u)\, du\bigg)\, \mathbf{e}(\ell z) \\ &=\; -\sum_{\ell \in \mathbb{Z}} \frac{\operatorname{sgn}(\ell - c) \,+\, \operatorname{sgn}(z_2)}{2}\, \mathbf{e}(\ell z) \;=\; \mathcal{R}_c(\mathbf{e}(z)) \end{aligned}$$

if $z_2 \neq 0$, and the formula remains true for $z_2 = 0$ by continuity. An alternative proof can be obtained by noting that $\mathbf{e}(-cz)\mathcal{R}_c(\mathbf{e}(z))$ is periodic of period 1 with respect to c and expanding it as a Fourier series in c, again by the Poisson summation formula.

For $m \in \mathbb{N}$ and $s = (\alpha, \beta) \in \mathbb{Q}^2$ we now define a *universal Appell-Lerch sum* $\mathcal{A}^s_m(\tau, z)$ by

$$\mathcal{A}^s_m(\tau, z) \;=\; \mathbf{e}(-m\alpha z_s)\, \operatorname{Av}^{(m)}\big[\mathcal{R}_{-2m\alpha}(y/y_s)\big] \qquad (y_s = \mathbf{e}(z_s) = \mathbf{e}(\beta) q^\alpha)\,. \tag{8.17}$$

It is easy to check that this function satisfies

$$\mathcal{A}_m^{(\alpha+\lambda,\beta+\mu)} \;=\; \mathbf{e}(-m(\mu\alpha-\lambda\beta))\,\mathcal{A}_m^{(\alpha,\beta)} \qquad (\lambda,\,\mu\in\mathbb{Z}) \tag{8.18}$$

and hence, in view of the corresponding property (8.7) of D_s, that the product $D_s(\tau)\mathcal{A}_m^s(\tau,z)$ depends only on the class of s in $S/\mathbb{Z}^2$. We can therefore define

$$\varphi^P(\tau,z) \;:=\; \sum_{s\in S/\mathbb{Z}^2} D_s(\tau)\,\mathcal{A}_m^s(\tau,z)\,, \tag{8.19}$$

and it is obvious from the above discussion that this function, which we will call the **polar part** of φ, satisfies the index m elliptic transformation property (4.2) and has the same poles and residues as φ, so that the difference $\varphi-\varphi^P$ is holomorphic and has a theta expansion. In fact, we have:

Theorem 8.1. *Let $\varphi(\tau,z)$ be a meromorphic Jacobi form with simple poles at $z=z_s=\alpha\tau+\beta$ for $s=(\alpha,\beta)\in S\subset\mathbb{Q}^2$, with Fourier coefficients $h_\ell(\tau)$ defined by (8.1) and (8.2) or by (8.4) and residues $D_s(\tau)$ defined by (8.6). Then φ has the decomposition*

$$\varphi(\tau,z) \;=\; \varphi^F(\tau,z) \,+\, \varphi^P(\tau,z)\,, \tag{8.20}$$

where φ^F and φ^P are defined by equations (8.3) and (8.19), respectively.

Proof: Fix a point $P=A\tau+B\in\mathbb{C}$ with $(A,B)\in\mathbb{R}^2\smallsetminus S$. Since φ, φ^F and φ^P are meromorphic, it suffices to prove the decomposition (8.20) on the horizontal line $\operatorname{Im}(z)=\operatorname{Im}(P)=A\tau_2$. On this line we have the Fourier expansion

$$\varphi(\tau,z) \;=\; \sum_{\ell\in\mathbb{Z}} q^{\ell^2/4m}\,h_\ell^{(P)}(\tau)\,y^\ell\,,$$

where the coefficients $h_\ell^{(P)}$ are defined by (8.1) (modified as explained in the text there if $A=\alpha$ for any $(\alpha,\beta)\in S$, but for simplicity we will simply assume that this is not the case, since we are free to choose A any way we want). Comparing this with (8.3) gives

$$\varphi(\tau,z) \,-\, \varphi^F(\tau,z) \;=\; \sum_{\ell\in\mathbb{Z}}\bigl(h_\ell^{(P)}(\tau)-h_\ell(\tau)\bigr)\,q^{\ell^2/4m}\,y^\ell \qquad (\operatorname{Im}(z)=\operatorname{Im}(P))\,. \tag{8.21}$$

But $q^{\ell^2/4m}(h_\ell^{(P)}(\tau)-h_\ell(\tau))$ is just $2\pi i$ times the sum of the residues of $\varphi(\tau,z)\mathbf{e}(-\ell z)$ in the parallelogram with width 1 and horizontal sides at heights $A\tau_2$ and $-\ell\tau_2/2m$, with the residues of any poles on the latter line being counted with multiplicity 1/2 because of the way we defined h_ℓ in that case, so

$$\begin{aligned} q^{\ell^2/4m}\bigl(h_\ell^{(P)}(\tau)-h_\ell(\tau)\bigr) \;&=\; 2\pi i \sum_{s=(\alpha,\beta)\,\in\, S/\mathbb{Z}} \frac{\operatorname{sgn}(\alpha-A)-\operatorname{sgn}(\alpha+\ell/2m)}{2}\,\operatorname{Res}_{z=z_s}\bigl(\varphi(\tau,z)\mathbf{e}(-\ell z)\bigr)\\ &=\; \sum_{s=(\alpha,\beta)\,\in\, S/\mathbb{Z}} \frac{\operatorname{sgn}(\alpha-A)-\operatorname{sgn}(\ell+2m\alpha)}{2}\;D_s(\tau)\,\mathbf{e}(-(\ell+m\alpha)z_s)\,. \end{aligned}$$

(Here "$(\alpha,\beta)\in S/\mathbb{Z}$" means that we consider all α, but β only modulo 1, which is the same by periodicity as considering only the (α,β) with $B\le\beta<B+1$.) Inserting this formula into (8.21) and using (8.15), we find

$$\begin{aligned}\varphi(\tau,z)\;-\;\varphi^F(\tau,z)\;&=\;-\sum_{s=(\alpha,\beta)\,\in\,S/\mathbb{Z}}\mathbf{e}(-m\alpha z_s)\,D_s(\tau)\sum_{\ell\in\mathbb{Z}}\frac{\operatorname{sgn}(\operatorname{Im}(z-z_s))+\operatorname{sgn}(\ell+2m\alpha)}{2}\,\Bigl(\frac{y}{y_s}\Bigr)^{\ell}\\&=\;\sum_{s=(\alpha,\beta)\,\in\,S/\mathbb{Z}}\mathbf{e}(-m\alpha z_s)\;D_s(\tau)\;\mathcal{R}_{-2m\alpha}(y/y_s)\\&=\;\sum_{s=(\alpha,\beta)\,\in\,S/\mathbb{Z}^2}\;\sum_{\lambda\in\mathbb{Z}}\mathbf{e}\bigl(-m(\alpha-\lambda)(z_s-\lambda\tau)\bigr)\,D_{(\alpha-\lambda,\beta)}(\tau)\;\mathcal{R}_{-2m(\alpha-\lambda)}(q^{\lambda}y/y_s)\\&=\;\sum_{s=(\alpha,\beta)\,\in\,S/\mathbb{Z}^2}D_s(\tau)\,\mathbf{e}(-m\alpha z_s)\;\sum_{\lambda\in\mathbb{Z}}q^{m\lambda^2}\,y^{2m\lambda}\,\mathcal{R}_{-2m\alpha}(q^{\lambda}y/y_s)\,,\end{aligned}$$

where in the last line we have used the periodicity property (8.7) of $D_s(\tau)$ together with the obvious periodicity property $\mathcal{R}_{c+n}(y)=y^n\mathcal{R}_c(y)$ of $\mathcal{R}_c(y)$. But the inner sum in the last expression is just $\mathrm{Av}^{(m)}\bigl[\mathcal{R}_{-2m\alpha}(y/y_s)\bigr]$, so from the definition (8.17) we see that this agrees with $\varphi^P(\tau,z)$, as claimed.

8.3 Mock modularity of the Fourier coefficients

In subsections §8.1 and §8.2 we introduced a canonical splitting of a meromorphic Jacobi form φ into a finite part φ^F and a polar part φ^P, but there is no reason yet (apart from the simplicity of equation (8.2)) to believe that the choice we have made is the "right" one: we could have defined periodic Fourier coefficients $h_\ell(\tau)$ in many other ways (for instance, by taking $P=P_0-\ell/2m\tau$ with any fixed $P_0\in\mathbb{C}$ or more generally $P=P_\ell-\ell\tau/2m$ where P_ℓ depends only on ℓ modulo $2m$) and obtained other functions φ^F and φ^P. What makes the chosen decomposition special is that, as we will now show, the Fourier coefficients defined in (8.2) are (mixed) mock modular forms and the function φ^F therefore a (mixed) mock Jacobi form in the sense of §7.3. The corresponding shadows will involve theta series that we now introduce.

For $m\in\mathbb{N}$, $\ell\in\mathbb{Z}/2m\mathbb{Z}$ and $s=(\alpha,\beta)\in\mathbb{Q}^2$ we define the unary theta series

$$\Theta^s_{m,\ell}(\tau)\;=\;\mathbf{e}(-m\alpha\beta)\sum_{\lambda\,\in\,\mathbb{Z}+\alpha+\ell/2m}\lambda\,\mathbf{e}(2m\beta\lambda)\,q^{m\lambda^2}\tag{8.22}$$

of weight 3/2 and its Eichler integral[17]

$$\Theta^{s\,*}_{m,\ell}(\tau)\;=\;\frac{\mathbf{e}(m\alpha\beta)}{2}\sum_{\lambda\,\in\,\mathbb{Z}+\alpha+\ell/2m}\operatorname{sgn}(\lambda)\,\mathbf{e}(-2m\beta\lambda)\operatorname{erfc}\bigl(2|\lambda|\sqrt{\pi m\tau_2}\bigr)\,q^{-m\lambda^2}\tag{8.23}$$

---

[17] Strictly speaking, the Eichler integral as defined by equation (7.2) with $k=1/2$ would be this multiplied by $2\sqrt{\pi/m}$, but this normalization will lead to simpler formulas and, as already mentioned, there is no good universal normalization for the shadows of mock modular forms.

(cf. (7.2) and (7.6)). One checks easily that these functions transform by

$$\begin{aligned}\Theta^{(\alpha+\lambda,\beta+\mu)}_{m,\ell}(\tau) &= \mathbf{e}(m(\mu\alpha-\lambda\beta))\,\Theta^{(\alpha,\beta)}_{m,\ell}(\tau) \qquad (\lambda,\,\mu\in\mathbb{Z})\,, & (8.24)\\ \Theta^{(\alpha+\lambda,\beta+\mu)\,*}_{m,\ell}(\tau) &= \mathbf{e}(-m(\mu\alpha-\lambda\beta))\,\Theta^{(\alpha,\beta)\,*}_{m,\ell}(\tau) \qquad (\lambda,\,\mu\in\mathbb{Z})\,. & (8.25)\end{aligned}$$

with respect to translations of s by elements of $\mathbb{Z}^2$. From this and (8.7) it follows that the products $D_s\overline{\Theta^s_{m,\ell}}$ and $D_s\Theta^{s\,*}_{m,\ell}$ depend only on the class of s in $S/\mathbb{Z}^2$, so that the sums over s occurring in the following theorem make sense.

Theorem 8.2. *Let φ, h_ℓ and φ^F be as in Theorem 8.1. Then each h_ℓ is a mixed mock modular form of weight $k-\frac{1}{2}\,|\,k-1$, with shadow $\sum_{s\in S/\mathbb{Z}^2} D_s(\tau)\,\overline{\Theta^s_{m,\ell}(\tau)}$, and the function φ^F is a mixed mock Jacobi form. More precisely, for each $\ell\in\mathbb{Z}/2m\mathbb{Z}$ the completion of h_ℓ defined by*

$$\widehat{h}_\ell(\tau) \;:=\; h_\ell(\tau) \;-\; \sum_{s\in S/\mathbb{Z}^2} D_s(\tau)\,\Theta^{s\,*}_{m,\ell}(\tau)\,, \qquad (8.26)$$

with $\Theta^{s\,*}_{m,\ell}$ as in (8.23), tranforms like a modular form of weight $k-1/2$ with respect to some congruence subgroup of $SL(2,\mathbb{Z})$, and the completion of φ^F defined by

$$\widehat{\varphi}^F(\tau,z) \;:=\; \sum_{\ell\ (\mathrm{mod}\ 2m)} \widehat{h}_\ell(\tau)\,\vartheta_{m,\ell}(\tau,z) \qquad (8.27)$$

transforms like a Jacobi form of weight k and index m with respect to the full modular group.

The key property needed to prove this theorem is the following proposition, essentially due to Zwegers, which says that the functions $\mathcal{A}^s_m(\tau,z)$ defined in §8.2 are (meromorphic) mock Jacobi forms of weight 1 and index m, with shadow $\sum_{\ell\ (\mathrm{mod}\ 2m)}\overline{\Theta^s_{m,\ell}(\tau)}\,\vartheta_{m,\ell}(\tau,z)$ (more precisely, that each $\mathcal{A}^s_m$ is a meromorphic mock Jacobi form of this weight, index and shadow with respect to some congruence subgroup of $SL(2,\mathbb{Z})$ depending on s and that the collection of all of them is a vector-valued mock Jacobi form on the full modular group):

Proposition 8.2. *For $m\in\mathbb{N}$ and $s\in\mathbb{Q}^2$ the completion $\widehat{\mathcal{A}}^s_m$ of $\mathcal{A}^s_m$ defined by*

$$\widehat{\mathcal{A}}^s_m(\tau,z) \;:=\; \mathcal{A}^s_m(\tau,z) \;+\; \sum_{\ell\ (\mathrm{mod}\ 2m)} \Theta^{s\,*}_{m,\ell}(\tau)\,\vartheta_{m,\ell}(\tau,z)\;. \qquad (8.28)$$

satisfies

$$\begin{aligned}\widehat{\mathcal{A}}^{(\alpha+\lambda,\beta+\mu)}_m(\tau,z) &= \mathbf{e}(-m(\mu\alpha-\lambda\beta))\,\widehat{\mathcal{A}}^{(\alpha,\beta)}_m(\tau) \qquad (\lambda,\,\mu\in\mathbb{Z})\,, & (8.29)\\ \widehat{\mathcal{A}}^s_m(\tau,z+\lambda\tau+\mu) &= \mathbf{e}(-m(\lambda^2\tau+2\lambda z))\,\widehat{\mathcal{A}}^s_m(\tau) \qquad (\lambda,\,\mu\in\mathbb{Z})\,, & (8.30)\\ \widehat{\mathcal{A}}^s_m\Big(\frac{a\tau+b}{c\tau+d},\,\frac{z}{c\tau+d}\Big) &= (c\tau+d)\,\mathbf{e}\Big(\frac{mcz^2}{c\tau+d}\Big)\,\widehat{\mathcal{A}}^{s\gamma}_m(\tau,z) \quad \Big(\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in SL(2,\mathbb{Z})\Big)\,. & (8.31)\end{aligned}$$

Proof: The first two properties are easy to check because they hold for each term in (8.28) separately. The modular transformation property, considerably less obvious, is essentially the content of Proposition 3.5 of [130], but the functions he studies are different from ours and we need a small calculation to relate them. Zwegers defines two functions $f_u^{(m)}(z;\tau)$ and $\tilde{f}_u^{(m)}(z;\tau)$ $(m \in \mathbb{N},\ \tau \in \mathbb{H},\ z,\, u \in \mathbb{C})$ by

$$f_u^{(m)}(z;\tau) = \mathrm{Av}^{(m)}\Big[\frac{1}{1-y/\mathbf{e}(u)}\Big], \qquad \tilde{f}_u^{(m)}(z;\tau) = f_u^{(m)}(z;\tau) - \frac{1}{2}\sum_{\ell \ (\mathrm{mod}\ 2m)} R_{m,\ell}(u;\tau)\,\vartheta_{m,\ell}(\tau,z)$$

(here we have rewritten Definition 3.2 of [130] in our notation), where

$$R_{m,\ell}(u;\tau) \;=\; \sum_{r\in \ell+2m\mathbb{Z}} \Big\{\mathrm{sgn}\big(r+\frac{1}{2}\big) \,-\, \mathrm{erf}\Big(\sqrt{\pi}\,\frac{r\tau_2+2mu_2}{\sqrt{m\tau_2}}\Big)\Big\}\, q^{-r^2/4m}\,\mathbf{e}(-ru)\,, \tag{8.32}$$

and shows (Proposition 3.5) that $\tilde{f}_u^{(m)}$ satisfies the modular transformation property

$$\tilde{f}^{(m)}_{u/(c\tau+d)}\Big(\frac{z}{c\tau+d};\frac{a\tau+b}{c\tau+d}\Big) \;=\; (c\tau+d)\,\mathbf{e}\Big(\frac{mc(z^2-u^2)}{c\tau+d}\Big)\,\tilde{f}_u^{(m)}(z;\tau) \tag{8.33}$$

for all $\gamma = \begin{pmatrix} a & b\\ c & d\end{pmatrix} \in SL(2,\mathbb{Z})$. Noting that $\mathrm{erf}(x) = \mathrm{sgn}(x)(1-\mathrm{erfc}(|x|))$, we find that

$$\frac{1}{2}\,R_{m,\ell}(z_s;\tau) \;=\; \sum_{r\equiv \ell \ (\mathrm{mod}\ 2m)} \frac{\mathrm{sgn}(r+\frac{1}{2})-\mathrm{sgn}(r+2m\alpha)}{2}\,q^{-r^2/4m}\,y_s^{-r} \;+\; \mathbf{e}(m\alpha z_s)\,\Theta^{s\,*}_{m,l}(\tau)$$

in our notation. On the other hand, from (8.13) we have

$$\mathcal{R}_{-2m\alpha}(y) \;=\; \frac{1}{y-1} \;+\; \sum_{r\in\mathbb{Z}} \frac{\mathrm{sgn}(r+\frac{1}{2})-\mathrm{sgn}(r+2m\alpha)}{2}\,y^r \tag{8.34}$$

(note that the sum here is finite). Replacing y by y/y_s and applying the operator $\mathrm{Av}^{(m)}$, we find (using (8.10))

$$\mathbf{e}(m\alpha z_s)\,\mathcal{A}^s_m(\tau,z) \;=\; -f^{(m)}_{z_s}(z;\tau) \;+\; \sum_{r\in\mathbb{Z}} \frac{\mathrm{sgn}(r+\frac{1}{2})-\mathrm{sgn}(r+2m\alpha)}{2}\,q^{-r^2/4m}\,y_s^{-r}\,\vartheta_{m,r}(\tau,z)\,.$$

Combining these two equations and rearranging, we obtain

$$\widehat{\mathcal{A}}^s_m(\tau,z) \;=\; -\,\mathbf{e}(-m\alpha z_s)\,\tilde{f}^{(m)}_{z_s}(z;\tau)\,,$$

and the modularity property (8.31) then follows from (8.33) after a short calculation.

The proof of Theorem 8.2 follows easily from Proposition 8.2. We define the completion of the function φ^P studied in §8.2 by

$$\widehat{\varphi}^P(\tau,z) \;:=\; \sum_{s\in S/\mathbb{Z}^2} D_s(\tau)\,\widehat{\mathcal{A}}^s_m(\tau,z)\,. \tag{8.35}$$

The sum makes sense by (8.29), and from the transformation equations (8.30)–(8.31) together with the corresponding properties (8.7)–(8.8) of the residue functions $D_s(\tau)$ it follows that $\widehat{\varphi}^P(\tau,z)$ transforms like a Jacobi form of weight k and index m with respect to the full modular group. Comparing equations (8.35) and (8.28) with equations (8.27) and (8.26), we find that

$$\widehat{\varphi}^P(\tau,z) \,-\, \varphi^P(\tau,z) \;=\; \sum_{\ell\,\in\,\mathbb{Z}/2m\mathbb{Z}}\;\sum_{s\,\in\,S/\mathbb{Z}^2} D_s(\tau)\,\Theta^{s\,*}_{m,\ell}(\tau)\,\vartheta_{m,\ell}(\tau,z) \;=\; \varphi^F(\tau,z) \,-\, \widehat{\varphi}^F(\tau,z)$$

and hence, using Theorem 8.1, that

$$\widehat{\varphi}^F(\tau,z) \,+\, \widehat{\varphi}^P(\tau,z) \;=\; \varphi^F(\tau,z) \,+\, \varphi^P(\tau,z) \;=\; \varphi(\tau,z)\,.$$

Since both $\varphi(\tau,z)$ and $\widehat{\varphi}^P(\tau,z)$ transform like Jacobi forms of weight k and index m, it follows that $\widehat{\varphi}^F(\tau,z)$ also does, and then the fact that each $\widehat{h}_\ell$ transforms like a modular form of weight $k-1/2$ (and hence that each h_ℓ is a mixed mock modular form with the weight and shadow given in the theorem) follows by the same argument that proves the modularity of the coefficients h_ℓ in the theta expansion (4.10) in the classical case.

Summary. For the reader's convenience, we give a brief summary of the results given up to now. We have the following six functions of two variables $(\tau,z)\in\mathbb{H}\times\mathbb{C}$:

- $\varphi(\tau,z)$, a meromorphic Jacobi form of weight k and index m, assumed to have only simple poles at $z=z_s=\alpha\tau+\beta$ for $s=(\alpha,\beta)$ in some discrete subset $S\subset\mathbb{Q}^2$;
- $\varphi^F(\tau,z)$, the **finite** part of φ, defined by the theta expansion $\sum_{\ell \pmod{2m}} h_\ell(\tau)\vartheta_{m,\ell}(\tau,z)$ where $h_\ell(\tau)$ is $q^{\ell^2/4m}$ times the ℓth Fourier coefficient of $\varphi(\tau,z-\ell\tau/2m)$ on the real line;
- $\varphi^P(\tau,z)$, the **polar** part of φ, defined as $\sum_{s\in S/\mathbb{Z}^2} D_s(\tau)\mathcal{A}^s_m(\tau,z)$, where $\mathcal{A}^s_m$ is an explicit Appell-Lerch sum having simple poles at $z\in z_s+\mathbb{Z}\tau+\mathbb{Z}$;
- $\varphi^C(\tau,z)$, a non-holomorphic **correction** term, defined as $\sum_s\sum_\ell D_s(\tau)\Theta^{s\,*}_{m,\ell}(\tau)\vartheta_{m,\ell}(\tau,z)$ where the $\Theta^{s\,*}_{m,\ell}$ are the Eichler integrals of explicit unary theta series of weight $3/2$;
- $\widehat{\varphi}^F(\tau,z)$, the completed finite part, defined as $\varphi^F(\tau,z)-\varphi^C(\tau,z)$;
- $\widehat{\varphi}^P(\tau,z)$, the completed polar part, defined as $\varphi^P(\tau,z)+\varphi^C(\tau,z)$.

These six functions are related by

$$\varphi^F \,+\, \varphi^P \;=\; \varphi \;=\; \widehat{\varphi}^F \,+\, \widehat{\varphi}^P\,, \qquad \varphi^F \,-\, \widehat{\varphi}^F \;=\; \varphi^C \;=\; \widehat{\varphi}_P \,-\, \varphi_P\,. \tag{8.36}$$

Each of them is real-analytic in τ, meromorphic in z, satisfies the elliptic transformation property (4.2) with respect to z, and has precisely two of four further desirable properties of such a function (note that $6 = \binom{4}{2}$), as shown in the following table

	φ	φ^F	φ^{P}	φ^C	$\widehat{\varphi}^{\mathrm{F}}$	$\widehat{\varphi}^{\mathrm{P}}$
holomorphic in τ ?	✓	✓	✓	–	–	–
transforms like a Jacobi form ?	✓	–	–	–	✓	✓
holomorphic in z ?	–	✓	–	✓	✓	–
determined by the poles of φ ?	–	–	✓	✓	–	✓

in which the three checked entries in each row correspond to one of the equations (10.1). Each Fourier coefficient h_ℓ of φ is a mixed mock modular form of weight $(k-1, 1/2)$, and the finite part φ^F is a mixed mock Jacobi form. In the holomorphic case, the functions φ, φ^F and $\widehat{\varphi}^F$ coincide and the functions φ^P, φ^C and $\widehat{\varphi}^P$ vanish.

We end by mentioning one further property of the canonical decomposition $\varphi = \varphi^F + \varphi^P$ that seems of interest. The finite part φ^P of φ is a linear combination of terms $h_\ell(\tau)\,\vartheta_{m,\ell}(\tau, z)$ where $h_\ell(\tau)$ is a Laurent power series in q and $\vartheta_{m,\ell}(\tau, z)$ a linear combination of terms $q^{r^2/4m}y^r$, so it satisfies the same growth condition as a weak Jacobi form, viz., that it contains only monomials $q^n y^r$ with discriminant $4mn - r^2$ bounded from below. But the Appell-Lerch sums $\mathcal{A}_m^s(\tau, z)$ occurring in $\varphi^P(\tau, z)$ have the opposite property: substituting the Fourier expansion (8.15) of $\mathcal{R}_c$ into the definition (8.17), we find that they contain only terms $q^n y^r$ with $4mn - r^2 = -\delta^2$ where the numbers δ occurring are rational (more precisely, $\delta \in r + 2m(\mathbb{Z} + \alpha)$) and unbounded (more precisely, δ lies between 0 and $r + \mathrm{O}(1)$).

8.4 The case of double poles

In this subsection we extend our considerations to the case when φ is allowed to have double poles, again assumed to be at points $z = z_s = \alpha\tau + \beta$ for $s = (\alpha, \beta)$ belonging to some discrete subset S of $\mathbb{Q}^2$. The first thing we need to do is to generalize the definition (8.6) to this case. For $s \in S$ we define functions E_s and D_s on $\mathbb{H}$ by the Laurent expansion

$$\mathbf{e}(m\alpha z_s)\,\varphi(\tau, z_s + \varepsilon) \;=\; \frac{E_s(\tau)}{(2\pi i\varepsilon)^2} + \frac{D_s(\tau) - 2m\alpha\, E_s(\tau)}{2\pi i\varepsilon} + O(1) \qquad \text{as } \varepsilon \to 0. \tag{8.37}$$

(Notice that $D_s(\tau)$ is the same function as in (8.6) when the pole is simple.) It is easily checked that the behavior of these functions under translations of s by elements of $\mathbb{Z}^2$ is given by equation (8.7) and its analogue for E_s. For the modular behavior, we have:

Proposition 8.3. *The functions $E_s(\tau)$ and $D_s(\tau)$ defined by (8.37) are modular forms of weight $k-2$ and $k-1$, respectively. More respectively, for all $s \in S$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2, \mathbb{Z})$ we have*

$$E_s\Big(\frac{a\tau + b}{c\tau + d}\Big) \;=\; (c\tau + d)^{k-2} E_{s\gamma}(\tau)\,, \qquad D_s\Big(\frac{a\tau + b}{c\tau + d}\Big) \;=\; (c\tau + d)^{k-1} D_{s\gamma}(\tau)\,. \tag{8.38}$$

Proof: We rewrite (8.37) as

$$\mathbf{e}(m\alpha(z_s+2\varepsilon))\,\varphi(\tau,z_s+\varepsilon) \;=\; \frac{E_s(\tau)}{(2\pi i\varepsilon)^2} \,+\, \frac{D_s(\tau)}{2\pi i\varepsilon} \,+\, \mathrm{O}(1)\,,$$

and also write α_s and $z_s(\tau)$ instead of just α and z_s. Then using the easily checked identities

$$z_s(\gamma\tau) \;=\; \frac{z_{s\gamma}(\tau)}{c\tau+d}\,, \qquad \alpha_{s\gamma}\,z_{s\gamma}(\tau) \,-\, \alpha_s\,z_s(\gamma\tau) \;=\; \frac{c\,z_{s\gamma}(\tau)^2}{c\tau+d} \qquad \Bigl(\gamma\tau := \frac{a\tau+b}{c\tau+d}\Bigr)\,,$$

and the modular transformation equation (4.1), we find

$$\begin{aligned}
\frac{(c\tau+d)^2\,E_s(\gamma\tau)}{(2\pi i\varepsilon)^2} + \frac{(c\tau+d)\,D_s(\gamma\tau)}{2\pi i\varepsilon} &\equiv \mathbf{e}\Bigl(m\alpha\Bigl(z_s(\gamma\tau)+\frac{2\varepsilon}{c\tau+d}\Bigr)\Bigr)\,\varphi\Bigl(\gamma\tau,\,z_s(\gamma\tau)+\frac{\varepsilon}{c\tau+d}\Bigr)\\
&\equiv \mathbf{e}\Bigl(m\alpha\frac{z_{s\gamma}(\tau)+2\varepsilon}{c\tau+d}\Bigr)\,\varphi\Bigl(\gamma\tau,\,\frac{z_{s\gamma}(\tau)+\varepsilon}{c\tau+d}\Bigr)\\
&\equiv (c\tau+d)^k\,\mathbf{e}\Bigl(m\alpha\frac{z_{s\gamma}+2\varepsilon}{c\tau+d}+\frac{mc(z_{s\gamma}+\varepsilon)^2}{c\tau+d}\Bigr)\,\varphi\bigl(\tau,\,z_{s\gamma}+\varepsilon\bigr)\\
&\equiv (c\tau+d)^k\,\mathbf{e}(m\alpha_{s\gamma}(z_{s\gamma}+2\varepsilon))\,\varphi(\tau,\,z_{s\gamma}+\varepsilon)\\
&\equiv (c\tau+d)^k\left[\frac{E_{s\gamma}(\tau)}{(2\pi i\varepsilon)^2}+\frac{D_{s\gamma}(\tau)}{2\pi i\varepsilon}\right],
\end{aligned}$$

where "$\equiv$" means equality modulo a quantity that is bounded as $\varepsilon\to 0$. The claim follows.

Next, we must define a standard Appell-Lerch sum with a double pole at $z=z_s$. We begin by defining a rational function $\mathcal{R}_c^{(2)}(y)$ with a double pole at $y=1$ for each $c\in\mathbb{R}$. Motivated by Proposition 8.1, we require

$$\mathcal{R}_c^{(2)}(\mathbf{e}(z)) \;=\; \mathrm{Av}_{\mathbb{Z}}\Bigl[\frac{\mathbf{e}(cz)}{(2\pi iz)^2}\Bigr] \;=\; \frac{1}{(2\pi i)^2}\sum_{n\in\mathbb{Z}}\frac{\mathbf{e}(c(z-n))}{(z-n)^2}\,. \tag{8.39}$$

To calculate this explicitly as a rational function, we could imitate the proof of Proposition 8.1, but it is easier to note that $\mathcal{R}_c^{(2)} = \bigl(-\frac{1}{2\pi i}\frac{d}{dz}+c\bigr)\mathcal{R}_c = \bigl(-y\frac{d}{dy}+c\bigr)\mathcal{R}_c$ and hence from equations (8.15), (8.14) and (8.13) we get three alternative definitions

$$\mathcal{R}_c^{(2)}(y) \;=\; \sum_{\ell\in\mathbb{Z}}\frac{|\ell-c|\,+\,\mathrm{sgn}(z_2)\,(\ell-c)}{2}\,y^\ell \tag{8.40}$$

$$= \begin{cases}\sum_{\ell\ge c}(\ell-c)\,y^\ell & \text{if } |y|<1\\ \sum_{\ell\le c}(c-\ell)\,y^\ell & \text{if } |y|>1\end{cases} \tag{8.41}$$

$$= y^{\lfloor c\rfloor+1}\Bigl(\frac{1}{(y-1)^2}+\frac{c\,-\,\lfloor c\rfloor}{y-1}\Bigr)\,. \tag{8.42}$$

of $\mathcal{R}^{(2)}_c(y)$. (Notice that in these equations neither the asterisk on the summation sign nor the case distinction for $c \in \mathbb{Z}$ and $c \notin \mathbb{Z}$ are needed as before, and that the function $\mathcal{R}^{(2)}_c(y)$, unlike $\mathcal{R}_c(y)$, is continuous in c.) For $s = (\alpha, \beta) \in \mathbb{Q}^2$ and $m \in \mathbb{N}$ we set

$$\mathcal{B}^s_m(\tau, z) \;=\; \mathbf{e}(-m\alpha z_s)\,\mathrm{Av}^{(m)}\bigl[\mathcal{R}^{(2)}_{-2m\alpha}(y/y_s)\bigr]\,, \tag{8.43}$$

in analogy with (8.17). If we then define the polar part φ^P of φ by

$$\varphi^P(\tau, z) \;=\; \sum_{s \in S/\mathbb{Z}^2} \Bigl(D_s(\tau)\,\mathcal{A}^s_m(\tau, z) \;+\; E_s(\tau)\,\mathcal{B}^s_m(\tau, z)\Bigr)\,; \tag{8.44}$$

then the definitions of the functions D_s, E_s, $\mathcal{A}^s_m$ and $\mathcal{B}^s_m$ immediately imply that φ^P has the same singularities as φ, so that the difference

$$\varphi^F(\tau, z) \;=\; \varphi(\tau, z) \;-\; \varphi^P(\tau, z) \tag{8.45}$$

is a holomorphic function of z.

As before, the key property of the Appell-Lerch sums is that they are again mock Jacobi forms, of a somewhat more complicated type than before. We introduce the unary theta series

$$\vartheta^s_{m,\ell}(\tau) \;=\; \mathbf{e}(-m\alpha\beta) \sum_{\lambda \in \mathbb{Z}+\alpha+\ell/2m} \mathbf{e}(2m\beta\lambda)\, q^{m\lambda^2} \tag{8.46}$$

of weight $1/2$ and its (again slightly renormalized) Eichler integral

$$\vartheta^{s\,*}_{m,\ell}(\tau) \;=\; \frac{\overline{\vartheta^s_{m,\ell}(\tau)}}{2\pi\sqrt{m\tau_2}} \;-\; \mathbf{e}(m\alpha\beta) \sum_{\lambda \in \mathbb{Z}+\alpha+\ell/2m} |\lambda|\,\mathbf{e}(-2m\beta\lambda)\,\mathrm{erfc}\bigl(2|\lambda|\sqrt{\pi m\tau_2}\bigr)\, q^{-m\lambda^2} \tag{8.47}$$

(cf. (7.2) and (7.6)). Then we can define the completion $\widehat{\mathcal{B}}^s_m$ of $\mathcal{B}^s_m$ by

$$\widehat{\mathcal{B}}^s_m(\tau, z) \;:=\; \mathcal{B}^s_m(\tau, z) \;+\; m \sum_{\ell \pmod{2m}} \vartheta^{s\,*}_{m,\ell}(\tau)\,\vartheta_{m,\ell}(\tau, z)\;. \tag{8.48}$$

Proposition 8.4. *For $m \in \mathbb{N}$ and $s \in \mathbb{Q}^2$ the completion $\widehat{\mathcal{B}}^s_m$ of $\mathcal{B}^s_m$ defined by (8.48) satisfies*

$$\widehat{\mathcal{B}}^{(\alpha+\lambda,\beta+\mu)}_m(\tau, z) \;=\; \mathbf{e}(-m(\mu\alpha - \lambda\beta + \lambda\mu))\,\widehat{\mathcal{B}}^{(\alpha,\beta)}_m(\tau) \qquad (\lambda,\,\mu \in \mathbb{Z})\,, \tag{8.49}$$

$$\widehat{\mathcal{B}}^s_m(\tau, z + \lambda\tau + \mu) \;=\; \mathbf{e}(-m(\lambda^2\tau + 2\lambda z))\,\widehat{\mathcal{B}}^s_m(\tau) \qquad (\lambda,\,\mu \in \mathbb{Z})\,, \tag{8.50}$$

$$\widehat{\mathcal{B}}^s_m\Bigl(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d}\Bigr) \;=\; (c\tau + d)^2\,\mathbf{e}\Bigl(\frac{mcz^2}{c\tau + d}\Bigr)\,\widehat{\mathcal{B}}^{s\gamma}_m(\tau, z) \qquad \Bigl(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z})\Bigr)\,. \tag{8.51}$$

The proof is exactly similar to that of 8.2. We define functions $g_u^{(m)}(z;\tau)$ and $\tilde{g}_u^{(m)}(z;\tau)$ by applying the operator $\dfrac{1}{2\pi i}\dfrac{\partial}{\partial u} - 2m\dfrac{u_2}{\tau_2}$ to $f_u^{(m)}(z;\tau)$ and $\tilde{f}_u^{(m)}(z;\tau)$; then the transformation equation (8.33) of $\tilde{f}_u^{(m)}$ implies the same transformation equation for $\tilde{g}_u^{(m)}$ but with the initial factor $(c\tau+d)$ replaced by $(c\tau+d)^2$, and a calculation exactly similar to the one given before shows that $\mathbf{e}(m\alpha z_s)\mathcal{B}_m^s(\tau,z)$ differs from $g_{z_s}^{(m)}(z;\tau)$ by a finite linear combination of functions $\vartheta_{m,r}(\tau,z)$ and that $\mathbf{e}(m\alpha z_s)\widehat{\mathcal{B}}_m^s(\tau,z) = \tilde{g}_{z_s}^{(m)}(z;\tau)$. We omit the details.

Theorem 8.3. *Let φ be as above, with singularities at $z = z_s$ ($s \in S \subset \mathbb{Q}^2$) given by (8.37). Then the finite part φ^F as defined by (8.45) coincides with the finite part defined by the theta expansion (8.3), the coefficients $h_\ell(\tau)$ in this expansion are mixed mock modular forms, with completion given by*

$$\widehat{h}_\ell(\tau) \;=\; h_\ell(\tau) \;+\; \sum_{s\in S/\mathbb{Z}^2}\Big(D_s(\tau)\,\Theta^{s\,*}_{m,\ell}(\tau) \;+\; E_s(\tau)\,\vartheta^{s\,*}_{m,\ell}(\tau)\Big)\,,$$

and the completion $\widehat{\varphi}^F$ defined by (8.27) transforms like a Jacobi form of weight k and index m.

The proof follows the same lines as before: the equivalence of (8.3) and (8.45) is proved by expanding $\varphi(\tau,z)$ as a Fourier series along the horizontal line $\mathrm{Im}(z) = \mathrm{Im}(P)$ for some generic point $P \in \mathbb{C}$ and calculating the difference $\varphi - \varphi^F$ as a sum of residues, and the mock modularity is proved by decomposing φ as $\widehat{\varphi}^F + \widehat{\varphi}^P$ with $\widehat{\varphi}^F$ as in (8.27) and $\widehat{\varphi}^P = \sum_{s\in S/\mathbb{Z}^2}\big(D_s\widehat{\mathcal{A}}_m^s + E_s\widehat{\mathcal{B}}_m^s\big)$, which transforms like a Jacobi form by virtue of Proposition 8.4. Again the details are left to the reader. Note that here the mock modular forms $h_\ell(\tau)$ are of the "even more mixed" variety mentioned at the end of §7.2, since they now have a shadow that is a linear combination of two terms $\sum_s D_s\,\overline{\Theta^s_{m,l}}$ and $\sum_s E_s\,\overline{\vartheta^s_{m,l}}$ belonging to two different tensor products $M_{k-1}\otimes\overline{M_{3/2}}$ and $M_{k-2}\otimes\overline{M_{1/2}}$ and hence two different mixed weights $k-\frac{1}{2}\,|\,k-1$ and $k-\frac{1}{2}\,|\,k-2$ rather than a single mixed weight $k-\frac{1}{2}\,|\,k-1$ as in the case of simple poles.

8.5 Examples

We end this section by giving five examples of meromorphic Jacobi forms and their canonical decompositions into a mock Jacobi form and a finite linear combination of Appell-Lerch sums. We systematically use the notations $A = \varphi_{-2,1}$, $B = \varphi_{0,1}$, $C = \varphi_{-1,2}$ for the three basic generators of the ring of weak Jacobi forms as described in equations (4.26)–(4.35).

Example 1: Weight 1, index 1, simple pole at $z = 0$. As our first example we take the Jacobi form $\varphi = C/A \in J^{\mathrm{mer}}_{1,1}$, which has a simple pole of residue $1/\pi i$ at $z = 0$ and a Fourier expansion beginning

$$\frac{y+1}{y-1} \,-\, (y^2-y^{-2})\,q \,-\, 2(y^3-y^{-3})\,q^2 \,-\, 2(y^4-y^{-4})\,q^3 \,-\, (2y^5+y^4-y^{-4}-2y^{-5})\,q^4 \,-\, \cdots\,.$$

Calculating the Fourier expansion of the polar part $\varphi^P = \mathrm{Av}^{(1)}\bigl[\frac{y+1}{y-1}\bigr]$, we find that it begins the same way, and indeed, we must have $\varphi = \varphi^P$ because the Fourier coefficients h_ℓ all vanish (we have $h_{-\ell} = -h_\ell$ because the weight is odd and $h_{\ell+2} = h_\ell$ because the index is 1). So here there is no mock Jacobi form at all, but only the polar correction term given by the Appell-Lerch sum, a kind of a "Cheshire cat" example which is all smile and no cat.

Example 2: Weight 1, index 2, simple pole at $z = 0$. As a second example take $\varphi = BC/A \in J^{\mathrm{mer}}_{1,2}$. Here we find

$$\begin{aligned}\varphi^P \;&=\; 12\,\mathrm{Av}^{(2)}\Bigl[\frac{y+1}{y-1}\Bigr] \;=\; 12\,\frac{y+1}{y-1} \;-\; 12(y^4-y^{-4})q^2 \;+\; 24(y^5-y^{-5})\,q^3 \;+\;\cdots\\ \varphi^F \;&=\; (y-y^{-1}) \;-\; (y^3+45y-45y^{-1}-y^{-3})q \;+\; (45y^3-231y+231y^{-1}-45y^{-3})q^2 \;+\;\cdots\,,\end{aligned}$$

and by comparing the shadows and the first few Fourier coefficients (cf. eq. (7.32)), we see that φ^F is the negative of the mock Jacobi form $\mathcal{F}_2(\tau,z) = h^{(2)}(\tau)\,\vartheta_1(\tau,2z)$ discussed in Example 4 of §7.2 that is conjecturally related to representations of M_{24}. Using (4.31) and dividing by $\vartheta_1(\tau,2z)$, we can rewrite the identity $\varphi = \varphi^P + \varphi^F$ as

$$\eta(\tau)^{-3}\,\frac{\varphi_{0,1}(\tau,z)}{\varphi_{-2,1}(\tau,z)} \;=\; -\,\frac{12}{\vartheta_1(\tau,2z)}\,\mathrm{Av}^{(2)}\Bigl[\frac{1+y}{1-y}\Bigr] \;-\; h^{(2)}(\tau)\,. \qquad (8.52)$$

One can check that $\vartheta_1(\tau,2z)\mu(\tau,z) = \mathrm{Av}^{(2)}\bigl[\frac{1+y}{1-y}\bigr]$, so this agrees with the splitting (7.39) coming from the decomposition of the elliptic genus of $K3$ into superconformal characters.

Example 3: Weight 1, index 1/2, simple pole at $z = 0$. We can also multiply (7.39) by $\vartheta_1(\tau,z)$ and write it in the form

$$\frac{B}{\sqrt{A}} \;=\; -12\,\mathrm{Av}^{(\frac12,-)}\Bigl[\frac{\sqrt{y}}{1-y}\Bigr] \;-\; h^{(2)}(\tau)\,\vartheta_1(\tau,z)\,, \qquad (8.53)$$

where $\mathrm{Av}^{(\frac12,-)}[F(y)] := \sum_{n\in\mathbb{Z}}(-1)^n q^{n^2/2}\,y^n\,F(q^n y)$. This decomposition also fits into the scheme explained in this section, but slightly modified because the index here is half-integral.

Example 4: Weight 2, index 0, double pole at $z = 0$. Now take $\varphi = B/A$, which, as we saw in §4.3, is just a multiple of the Weierstrass $\wp$-function.This example does not quite fit in with the discussion given so far, since we did not allow $m = 0$ in the definition (8.9) of the averaging operator ($m < 0$ doesn't work at all, because the sum defining the average diverges, and $m = 0$ is less interesting since a form of index 0 is clearly determined up to a function of τ alone by its singularities, so that in our discussion we excluded that case too), but nevertheless it works. The decomposition $\varphi = \varphi^P + \varphi^F$ takes the form

$$\frac{\varphi_{0,1}(\tau,z)}{\varphi_{-2,1}(\tau,z)} \;=\; 12\,\mathrm{Av}^{(0)}\Bigl[\frac{y}{(1-y)^2}\Bigr] \;+\; E_2(\tau)\,, \qquad (8.54)$$

with the finite part being simply the quasimodular Eisenstein series $E_2(\tau)$, which is also a mock modular form. (It is automatically independent of z since the index is 0.)

Example 5: Weight 2, index 1, double pole at $z = 0$. The next example is a special case of an infinite family that will be treated in detail in §9. Take $\varphi = B^2/A \in J^{\mathrm{mer}}_{2,1}$. Then we find

$$\begin{aligned}\varphi \;&=\; \Big(\frac{144\,y}{(1-y)^2} + y + 22 + y^{-1}\Big) + \Big(22y^2 + 152y - 636 + 152y^{-1} + 22y^{-2}\Big)\,q \\ &\quad + \Big(145y^3 - 636y^2 + 3831y - 7544 + 3831y^{-1} - 636y^{-2} + 145y^{-3}\Big)\,q^2 + \cdots\,, \\ \varphi^{\mathrm{P}} \;&=\; \mathrm{Av}^{(1)}\Big[\frac{144\,y}{(1-y)^2}\Big] \;=\; \frac{144\,y}{(1-y)^2} + 0\,q + 144\,(y^3+y^{-3})q^2 + 288\,(y^4+y^{-4})q^3 + \cdots\end{aligned}$$

and hence that $\varphi^F = \varphi - \varphi^P = \sum_{4n-r^2\ge -1} C(4n-r^2)\,q^n\,y^r$ with the first few $C(\Delta)$ given by

Δ	-1	0	3	4	7	8	11	12	15
$C(\Delta)$	1	22	152	-636	3831	-7544	33224	-53392	191937

.

On the other hand, we have the weak Jacobi form $E_4(\tau)\varphi_{-2,1}(\tau,z) = \sum C^*(4n-r^2)q^ny^r$ with

Δ	-1	0	3	4	7	8	11	12	15
$C^*(\Delta)$	1	-2	248	-492	4119	-7256	33512	-53008	192513

,

We see that $C(\Delta)$ and $C^*(\Delta)$ are very close to each other (actually, asymptotically the same) and that their difference is precisely -288 times the Hurwitz-Kronecker class number $H(d)$. We thus have $\varphi^F = E_4A - 288\mathcal{H}$, and we have found $\mathcal{H}(\tau,z)$, the simplest of all mock Jacobi forms, as the finite part of the meromorphic Jacobi form $(E_4A^2 - B^2)/288A$. We remark that the numbers $-C^*(\Delta)$ have an interesting interpretation as "traces of singular moduli" [128].

Example 6: Weight -5, index 1, simple poles at the 2-torsion points. Finally, we give an example with more than one pole. Take $\varphi = A^3/C \in J_{1,-5}$. This function has three poles, all simple, at the three non-trivial 2-torsion points on the torus $\mathbb{C}/(\mathbb{Z}\tau+\mathbb{Z})$. The three corresponding modular forms, each of weight -6, are given by

$$D_{(0,\frac12)}(\tau) \;=\; 16\,\frac{\eta(2\tau)^{12}}{\eta(\tau)^{24}}\,, \quad D_{(\frac12,0)}(\tau) \;=\; -\frac{1}{4}\,\frac{\eta\big(\frac{\tau}{2}\big)^{12}}{\eta(\tau)^{24}}\,, \quad D_{(\frac12,\frac12)}(\tau) \;=\; \frac{i}{4}\,\frac{\eta\big(\frac{\tau+1}{2}\big)^{12}}{\eta(\tau)^{24}}\,,$$

and one finds

$$\begin{aligned}\varphi \;=\; \varphi^{\mathrm{P}} \;&=\; D_{(0,\frac12)}(\tau)\,\mathrm{Av}^{(1)}\Big[\frac{1}{2}\,\frac{y-1}{y+1}\Big] + q^{1/4}\,D_{(\frac12,0)}(\tau)\,\mathrm{Av}^{(1)}\Big[\frac{1}{2y}\,\frac{y+\sqrt{q}}{y-\sqrt{q}}\Big] \\ &\quad + q^{1/4}\,D_{(\frac12,\frac12)}(\tau)\,\mathrm{Av}^{(1)}\Big[\frac{1}{2y}\,\frac{y-\sqrt{q}}{y+\sqrt{q}}\Big]\,, \qquad \varphi^F \;\equiv\; 0\,,\end{aligned}$$

another "Cheshire cat" example (of necessity, for the same reason as in Example 1, since again $m = 1$ and k is odd).

9. Special mock Jacobi forms

We now turn to the study of certain families of meromorphic Jacobi forms of weights 1 or 2 and arbitrary positive index and of their related mock Jacobi forms, the weight 2 case being the one relevant for the application to black hole microstate counting. In this section, we introduce these families and formulate a nunmber of properties of them that were found experimentally and that we will prove in §10. In particular, it will turn out that all of these functions can be expressed, using the Hecke-like operators introduced in §4.4, in terms of a collection of mock Jacobi forms $\mathcal{Q}_m$ defined (though not quite uniquely: there is some choice involved starting at $m = 91$) for all square-free m, of weight either 1 or 2 depending on whether m has an odd or even number of prime factors. These functions, which seem to have particularly nice arithmetic properties, include all of the examples studied in §7, and several other special mock modular forms that have appeared in the literature.

9.1 The simplest meromorphic Jacobi forms

The two simplest situations for the theory described in §8 are when the meromorphic Jacobi form $\varphi = \varphi_{k,m}$ satisfies either

- $k = 1$ and $\varphi(\tau, z)$ has a simple pole at $z = 0$ and no other poles in $\mathbb{C}/(\mathbb{Z}\tau + \mathbb{Z})$ or
- $k = 2$ and $\varphi(\tau, z)$ has a double pole at $z = 0$ and no other poles in $\mathbb{C}/(\mathbb{Z}\tau + \mathbb{Z})$,

since in these two cases the modular forms $D_s(\tau)$ (resp. $E_s(\tau)$) defined in §8.2 (resp. §8.4) are simply constants and the canonical Fourier coefficients h_ℓ of φ are pure, rather than mixed, mock modular forms. If we normalize φ by $\varphi(\tau, z) \sim (2\pi i z)^{-1}$ in the first case and $\varphi(\tau, z) \sim (2\pi i z)^{-2}$ in the second case as $z \to 0$, then any two choices of φ differ by a (weakly) holomorphic Jacobi form whose Fourier coefficients are true modular forms and are therefore well understood. It therefore suffices in principle to make some specific choice of $\varphi_{1,m}$ and $\varphi_{2,m}$ for each m and to study their finite parts $\varphi^{\mathrm{F}}_{k,m}$, but, as we shall see, the problem of choosing a good representative φ is very interesting and has several subtleties. For any choice of $\varphi_{k,m}$, we will use the notation $\Phi_{k,m} := \varphi^{\mathrm{F}}_{k,m}$, with decorations like $\Phi^{\mathrm{stand}}_{k,m}$ or $\Phi^{\mathrm{opt}}_{k,m}$ corresponding to the choice of $\varphi_{k,m}$. The shadow of the mock Jacobi form $\Phi_{k,m}$ is independent of the choice of representative and is a multiple of $\sum_{\ell \pmod{2m}} \overline{\vartheta^{2-k}_{m,\ell}(\tau)}\, \vartheta_{m,\ell}(\tau, z)$, where $\vartheta^1_{m,\ell}$ and $\vartheta^0_{m,\ell}$ are defined as at the end of §4.2.

The Jacobi forms $\varphi_{2,m}$ are the ones related to the Fourier coefficients ψ_m of $\Phi_{10}(\Omega)^{-1}$ that appear in the application to black holes (eq. (1.1)). Indeed, since the principal part of the pole of $\psi_m(\tau, z)$ at $z = 0$ equals $p_{24}(m+1)\,\Delta(\tau)^{-1}\,(2\pi i z)^{-2}$, we can write $\Delta(\tau)\psi_m(\tau, z)$ as the sum of $p_{24}(m+1)\,\varphi_{2,m}$ and a weak Jacobi form. If we make the simplest choice for $\varphi_{2,m}$, namely

$$\varphi^{\mathrm{stand}}_{2,m} \;=\; \frac{B^{m+1}}{12^{m+1} A}\,, \tag{9.1}$$

where $A = \varphi_{-2,1}$, $B = \varphi_{0,1}$ are the special Jacobi forms introduced in §4 (which has the correct principal part because $A(\tau,z) \sim (2\pi iz)^2$ and $B(\tau,z) \to 12$ as $z \to 0$), then the first few values of $\Delta(\tau)\,\psi_m(\tau,z) - p_{24}(m+1)\,\varphi_{2,m}^{\text{stand}}(\tau,z)$ can be read off from (5.16). For instance, we have $p_{24}(3) = 3200$ and the fourth equation of (5.16) says that $\Delta\psi_2 = 3200\,\varphi_{2,2}^{\text{stand}} + \frac{16}{9}E_4AB + \frac{10}{27}E_6A^2$.

In the weight one case we can also make a simple standard choice of φ. It has to be divisible by C because the weight is odd, and since $C \sim 4\pi iz$ for $z \to 0$ we can choose

$$\varphi_{1,m}^{\text{stand}} \;=\; \frac{B^{m-1}\,C}{2\cdot 12^{m-1}\,A} \;=\; \frac{C}{2}\,\varphi_{2,m-2}^{\text{stand}}\,. \tag{9.2}$$

as our standard functions. However, we emphasize that (9.1) and (9.2) are only the starting points for our discussion and will be replaced by better choices as we go along.

We will apply the decomposition of §8 to the functions φ. The polar part is independent of the choice of φ and is given by $\varphi_{k,m}^{\text{P}} = \mathcal{A}_{k,m}$, where $\mathcal{A}_{k,m}$ is given (in the notation of equations (8.17) and (8.43)) by

$$\mathcal{A}_{1,m} \;=\; \mathcal{A}_{1,m}^{(0,0)} \;=\; \text{Av}^{(m)}\Big[\frac{1}{2}\,\frac{y+1}{y-1}\Big]\,, \qquad \mathcal{A}_{2,m} \;=\; \mathcal{A}_{2,m}^{(0,0)} \;=\; \text{Av}^{(m)}\Big[\frac{y}{(y-1)^2}\Big] \tag{9.3}$$

or—written out explicitly—by

$$\mathcal{A}_{1,m}(\tau,z) \;=\; -\frac{1}{2}\sum_{s\in\mathbb{Z}} q^{ms^2}\,y^{2ms}\,\frac{1+q^sy}{1-q^sy}\,, \qquad \mathcal{A}_{2,m}(\tau,z) \;=\; \sum_{s\in\mathbb{Z}} \frac{q^{ms^2+s}\,y^{2ms+1}}{(1-q^sy)^2}\,. \tag{9.4}$$

Also independent of the choice of φ is the correction term which must be added to the finite part and subtracted from the polar part of $\varphi_{2,m}$ in order to make these functions transform like Jacobi forms. In particular, for the double pole case, we find from (8.48) that the completion $\widehat{\Phi}_{2,m}$ of $\Phi_{2,m} = \varphi_{2,m}^{\text{F}}$ for any choice of $\varphi_{2,m}$ is given by

$$\widehat{\Phi}_{2,m}(\tau,z) \;=\; \Phi_{2,m}(\tau,z) \;-\; \sqrt{\frac{m}{4\pi}}\sum_{\ell\ (\text{mod}\ 2m)} \big(\vartheta^0_{m,\ell}\big)^*(\tau)\,\vartheta_{m,\ell}(\tau,z)\,, \tag{9.5}$$

where $\vartheta^0_{m,\ell}(\tau) = \vartheta_{m,\ell}(\tau,0)$ and $(\vartheta^0_{m,\ell})^*$ is normalized[18] as in equations (7.2) and (7.6).

The functions $\mathcal{A}_{1,m}$, $\mathcal{A}_{2,m}$ have wall-crossings, so their Fourier expansions depend on the value of $z_2 = \text{Im}(z)$. In the range $0 < z_2 < \tau_2$, or $|q| < |y| < 1$, they are given by

$$\mathcal{A}_{1,m} \;=\; \Big(-\sum_{s\geq 0}\sum_{\ell\geq 0}{}^{*} + \sum_{s<0}\sum_{\ell\leq 0}{}^{*}\Big)\, q^{ms^2+\ell s}\,y^{2ms+\ell}\,, \tag{9.6}$$

$$\mathcal{A}_{2,m} \;=\; \Big(\sum_{s\geq 0}\sum_{\ell\geq 0}{}^{*} - \sum_{s<0}\sum_{\ell\leq 0}{}^{*}\Big)\,\ell\, q^{ms^2+\ell s}\,y^{2ms+\ell}\,, \tag{9.7}$$

---

[18] Note that the function $\vartheta^{(0,0)\,*}_{m,\ell}$ as defined in (8.47) is equal to $1/\sqrt{4\pi m}$ times $(\vartheta^0_{m,\ell})^*$.

where the asterisk on the summation sign has the usual meaning (count the term $\ell = 0$ with multiplicity 1/2). Hence the coefficient of $q^n y^r$ is zero unless $\Delta = 4mn - r^2$ is the negative of the square of an integer congruent to r modulo $2m$, in which case it is 0, $\pm\frac{1}{2}$, ± 1 or ± 2 if $k = 1$ and is 0, $\pm\sqrt{-\Delta}$ or $\pm 2\sqrt{-\Delta}$ if $k = 2$. (Compare the remark at the end of §8.3.) If we expand $\mathcal{A}_{1,m}$ and $\mathcal{A}_{2,m}$ in Fourier series for z_2 in some other interval $(N, N+1)\tau_2$, then the formulas (9.6) and (9.7) are unchanged except for a shift in the inequalities defining the summation. In summary, the Fourier coefficients of the polar part of φ, although not quite like the coefficients of a usual Jacobi form—they do not only depend on Δ and on r mod $2m$, they are non-zero for infinitely many negative Δ, and there is wall-crossing—are extremely simple and have very slow growth (at most like $\sqrt{|\Delta|}$).

The interesting part of the Fourier expansion of $\varphi_{k,m}$ is therefore the mock part $\Phi_{k,m}$. Here the choice of $\varphi_{k,m}$ makes a big difference. For instance, Examples 1–4 of §8.5 say that the mock Jacobi forms $\Phi^{\text{stand}}_{k,m}$ corresponding to the standard choices (9.1) and (9.2) for small m are

$$\Phi^{\text{stand}}_{1,1} = 0, \quad \Phi^{\text{stand}}_{1,2} = \frac{1}{24}\mathcal{F}_2\,, \quad \Phi^{\text{stand}}_{2,0} = \frac{1}{12}E_2\,, \quad \Phi^{\text{stand}}_{2,1} = \frac{1}{144}E_4\,A - 2\,\mathcal{H}\,, \tag{9.8}$$

and we see that this choice is a good one in the first three cases (indeed, in these cases there *is* no choice!) but that, as we have already observed, a better choice of φ in the fourth case would have been the function $\varphi^{\text{opt}}_{2,1} = \varphi^{\text{stand}}_{2,1} - \frac{1}{144}E_4 A$, with mock part $\Phi^{\text{opt}}_{2,1} = -2\mathcal{H}$, because it has no terms $q^n y^r$ with discriminant $\Delta = 4n - r^2$ smaller than zero and (consequently) has Fourier coefficients that grow polynomially in Δ (actually, like $\Delta^{1/2+\varepsilon}$) rather than exponentially in $\sqrt{\Delta}$. Our next main objective, to which we devote the next subsection, is therefore to replace our initial choice of $\varphi_{k,m}$ as given by (9.1) or (9.2) by an optimal choice $\varphi^{\text{opt}}_{k,m}$, where "optimal" means "with the negative discriminants $d = 4mn - r^2$ that occur being as little negative as possible" or, equivalently, "with Fourier coefficients growing as slowly as possible." This choice turns out to have Fourier coefficients of very much smaller growth than those of our initial choice $\varphi^{\text{stand}}_{k,m}$ in all cases, and to be unique in many cases.

9.2 Choosing the function φ for weight 2 and small index

We now consider in more detail the meromorphic Jacobi forms of weight 2 and index m with a double pole at the origin introduced in §9.1. Any such form will have the form

$$\varphi_{2,m} = \frac{\psi_{0,m+1}}{12^{m+1}A}\,, \qquad \psi_{0,m+1} \in B^{m+1} + A\cdot\widetilde{J}_{2,m} \subset \widetilde{J}_{0,m+1}\,. \tag{9.9}$$

For each m, we want to choose a meromorphic Jacobi form $\varphi^{\text{opt}}_{2,m}$ differing from $\varphi^{\text{stand}}_{2,m}$ by a weak Jacobi form of weight 2 and index m whose coefficients grow as slowly as possible. We are interested both in the form of the corresponding $\psi_{0,m+1} = \psi^{\text{opt}}_{0,m+1}$ as a polynomial in A and B with modular form coefficients, and in the nature of the finite part $\Phi^{\text{opt}}_{2,m} = \varphi^{\text{opt}}_{2,m} - \mathcal{A}_{2,m}$ as a mock Jacobi form. In this subsection we will look at the optimal choice for $0 \le m \le 11$ to see what phenomena occur, while in the next subsection, we report on the calculations for higher values up to about 250 and formulate the general results.

For **m=0**, we have to choose

$$\varphi^{\rm opt}_{2,0} \;=\; \varphi^{\rm stand}_{2,0} \;=\; \frac{B}{12A}\,, \qquad \psi^{\rm opt}_{0,1} \;=\; B$$

because there are no weak Jacobi forms of weight 2 and index 0. The mock part $\Phi^{\rm opt}_{2,0}$ here is $E_2(\tau)/12$, as we have already seen (Example 3 of §8.5).

For **m=1** the only freedom is the addition of a multiple of E_4A and (as recalled in §9.1) the optimal choice is

$$\varphi^{\rm opt}_{2,1} \;=\; \varphi^{\rm stand}_{2,1} \;-\; \frac{E_4A}{144}\,, \qquad \psi^{\rm opt}_{0,2} \;=\; B^2 \;-\; E_4A^2\,, \tag{9.10}$$

because then $\Phi^{\rm opt}_{2,1} \;=\; -2\mathcal{H}$ has no terms q^ny^r with $4n - r^2 < 0$ and its Fourier coefficients $-2H(4n-r^2)$ have only polynomial growth as $4n - r^2 \to \infty$.

Now we look at **m=2**. Here we can modify $12^3\,\Phi^{\rm stand}_{2,2} = B^3/A - 12^3\mathcal{A}_{2,2}$ by adding to it any linear combination of the two weak Jacobi forms E_4AB and E_6A^2. Each of these three functions has Fourier coefficients of the form $c(n,r) = C_\varphi(8n-r^2)$ (because the (n,r)-coefficient of a Jacobi form of prime or prime power index m and even weight always depends only on the discriminant $4mn - r^2$), with the first few given by

Δ	-4	-1	0	4	7	8	12	15	16
$C(12^3\Phi^{\rm stand}_{2,2};\Delta)$	1	32	366	528	1056	-10144	-14088	92832	-181068
$C(E_4AB;\Delta)$	1	8	-18	144	2232	-4768	-22536	96264	-151884
$C(E_6A^2;\Delta)$	1	-4	6	-480	1956	-2944	-28488	96252	-139452

It follows from this table that the linear combination

$$\varphi^{\rm opt}_{2,2} \;=\; \varphi^{\rm stand}_{2,2} \;-\; \frac{1}{576}\,E_4AB \;+\; \frac{1}{864}E_6A^2 \;=\; \frac{B^3 \;-\; 3E_4A^2B \;+\; 2E_6A^3}{12^3\,A} \tag{9.11}$$

has a mock Jacobi part $\Phi^{\rm opt}_{2,2} = \varphi^{\rm opt}_{2,2} - \mathcal{A}_{2,2}$ with no negative terms and much smaller coefficients:

d	-4	-1	0	4	7	8	12	15	16
$C(\Phi^{\rm opt}_{2,2};\Delta)$	0	0	$1/4$	$-1/2$	-1	-1	-2	-2	$-5/2$

On inspection of these numbers and the next few values, we recognize them as simple linear combinations of Hurwitz class numbers, namely:

$$C(\Phi^{\rm opt}_{2,2};\Delta) = \begin{cases} -H(\Delta) & \text{if } \Delta \equiv 7 \pmod 8\,, \\ -H(\Delta) - 2H(\Delta/4) & \text{if } \Delta \equiv 0 \pmod 4\,, \end{cases}$$

which says that $\Phi^{\rm opt}_{2,2}$ is simply $-\mathcal{H}|V_2$, with $\mathcal{H}$ as in (7.25) and V_2 as in (4.37). (Note, by the way, the similarity between (9.11) and (4.33). Together they say that $\varphi^{\rm opt}_{2,2} = (C/2A)^2$.)

The cases **m=3**, **m=4**, **m=5**, **m=7**, **m=8**, **m=9** and **m=11** are similar to the case $m = 2$. For each of these values of m the coefficient of $q^n y^r$ in $\Phi^{\rm stand}_{2,m}$ depends only on $\Delta = 4nm - r^2$, the number of negative Δ for which this coefficient is non-zero equals the dimension of the space of weak Jacobi forms of weight 2 and index m, and we can subtract a unique weak Jacobi form from $\varphi^{\rm stand}_{2,m}$ to obtain a meromorphic Jacobi form $\varphi^{\rm opt}_{2,m}$ whose finite part is a strongly holomorphic mock Jacobi form with very small Fourier coefficients. For example, for $m = 3$ the optimal choice of $\varphi^{\rm opt}_{2,m}$ is $\varphi^{\rm opt}_{2,3} = \psi^{\rm opt}_{0,4}/12^4 A$ with

$$\psi^{\rm opt}_{0,4} \;=\; B^4 \,-\, 6E_4A^2B^2 \,+\, 8E_6A^3B \,-\, 3E_4^2A^4 \tag{9.12}$$

and the first few Fourier coefficients of $12^4\,\Phi^{\rm stand}_{2,3}$ and $\Phi^{\rm opt}_{2,3}$ are given by

Δ	-9	-4	-1	0	3	8	11	12	15
$12^4\,C(\Phi^{\rm stand}_{2,3};\Delta)$	1	42	687	5452	1104	8088	-1488	-139908	-19629
$C(\Phi^{\rm opt}_{2,3};\Delta)$	0	0	0	$1/3$	$-1/3$	-1	-1	$-4/3$	-2

Again we recognize the coefficients $C(\Phi^{\rm opt}_{2,3};\Delta)$ as $-H(\Delta)$ if $9\nmid\Delta$ and $-H(\Delta)-3H(\Delta/9)$ if $9\,|\,\Delta$, so that just as for $m = 2$ we find that the mock part of $\varphi^{\rm opt}_{2,3}$ equals $-\mathcal{F}_1|V_3$. The same thing also happens for $m = 5$, 7 and 11, suggesting that for prime values of m we can always choose

$$\Phi^{\rm opt}_{2,p} \;=\; -\,\mathcal{H}|V_p \qquad (p\ \text{prime}). \tag{9.13}$$

This will be proved in §10. For the prime-power values $m = 4$, 8 and 9 this formula no longer holds, but the coefficients of $\Phi^{\rm opt}_{2,m}$ are still simple linear combinations of Hurwitz-Kronecker class numbers and we find the expressions

$$\Phi^{\rm opt}_{2,4} = -\mathcal{H}|V_4 + 2\mathcal{H}|U_2\,, \quad \Phi^{\rm opt}_{2,8} = -\mathcal{H}|V_8 + 2\mathcal{H}|U_2V_2\,, \quad \Phi^{\rm opt}_{2,9} = -\mathcal{H}|V_9 + 3\mathcal{H}|U_3\,,$$

where U_t is the Hecke-like operator defined in §4.4 that multiplies the index of a Jacobi form (or mock Jacobi form) by t^2. This suggests that for prime power values of m we can choose

$$\Phi^{\rm opt}_{2,m} \;=\; -2\,\mathcal{H}|\,\mathcal{V}^{(1)}_{2,m} \qquad (m\ \text{a prime power}) \tag{9.14}$$

where $\mathcal{V}^{(1)}_{k,t}$ for $t \in \mathbb{N}$ is the Hecke-like operator from Jacobi (or weakly holomorphic Jacobi, or mock Jacobi) forms of weight k and index 1 to forms of weight k and index t defined in (4.38). Formula (9.14) will also be proved in §10.

The data for the values discussed so far ($0 \le m \le 11$, $m \ne 6$, 10) is summarized in the table below, which expresses the weak Jacobi form $\psi^{\rm opt}_{0,m+1} = 12^{m+1}A\,\varphi^{\rm opt}_{2,m}$ for each of these values of m as a polynomial in A and B. In this table we notice many regularities, e.g., the coefficients of $E_4A^2B^{m-2}$, $E_6A^3B^{m-3}$ and $E_4^2A^4B^{m-4}$ in $\psi^{\rm opt}_{0,m}$ are always given by $-\binom{m}{2}$, $2\binom{m}{3}$ and $-3\binom{m}{4}$, respectively, and more generally, if we write $\psi^{\rm opt}_{0,m}$ as $\sum_{i=0}^{m} f_{m,i}\, A^i\, B^{m-i}$ with $f_{m,i} = f_{m,i}(\tau) \in M_{2i}$, $f_{m,0} = 1$, then the constant terms $f_{m,i}(\infty)$ of the modular forms $f_{m,i}$ (which can be read off from the table by setting $E_4 = E_6 = 1$ and $D = 0$) are given in each case by $f_{m,i}(\infty) = (-1)^{i-1}(i-1)\binom{m}{i}$ or equivalently by the generating function

$$\sum_{i=0}^{m} f_{m,i}(\infty)\, X^{m-i} \;=\; (X-1)^m \,+\, m\,(X-1)^{m-1} \;=\; (X-1)^{m-1}(X+m-1)\,, \qquad (9.15)$$

e.g. $X^8 - 28X^6 + 112X^5 - 210X^4 + 224X^3 - 140X^2 + 48X - 7 \,=\, (X-1)^7(X+7)$. This formula has a simple interpretation: the constant term with respect to q of $\sum f_{m,i}\, A^i\, B^{m-i}$ is $\sum f_{m,i}(\infty)(y^{-1} - 2 + y)^i(y^{-1} + 10 + y)^{m-i}$ and this Laurent polynomial—subject to the condition $f_{m,1}(\infty) = 0$, which comes from the fact that there are no modular forms of weight 2 on $SL(2,\mathbb{Z})$—has minimal growth $\mathrm{O}\big(y + y^{-1}\big)$ precisely for the coefficients defined by (9.15).

m	1	2	3	4	5	6	8	9	10	12
B^m	1	1	1	1	1	1	1	1	1	1
$E_4 \cdot A^2B^{m-2}$	0	−1	−3	−6	−10	−15	−28	−36	−45	−66
$E_6 \cdot A^3B^{m-3}$	0	0	2	8	20	40	112	168	240	440
$E_4^2 \cdot A^4B^{m-4}$	0	0	0	−3	−15	−45	−210	−378	−630	−1485
$E_4E_6 \cdot A^5B^{m-5}$	0	0	0	0	4	24	224	504	1008	3168
$E_4^3 \cdot A^6B^{m-6}$	0	0	0	0	0	−5	−140	−420	−1050	−4620
$D \cdot A^6B^{m-6}$	0	0	0	0	0	1	10	30	48	222
$E_4^2E_6 \cdot A^7B^{m-7}$	0	0	0	0	0	0	48	216	720	4752
$E_4^4 \cdot A^8B^{m-8}$	0	0	0	0	0	0	−7	−63	−315	−3465
$E_4D \cdot A^8B^{m-8}$	0	0	0	0	0	0	−4	−36	−18	−360
$E_4^3E_6 \cdot A^9B^{m-9}$	0	0	0	0	0	0	0	8	80	1760
$E_6D \cdot A^9B^{m-9}$	0	0	0	0	0	0	0	20	−16	80
$E_4^5 \cdot A^{10}B^{m-10}$	0	0	0	0	0	0	0	0	−9	−594
$E_4^2D \cdot A^{10}B^{m-10}$	0	0	0	0	0	0	0	0	18	702
$E_4^4E_6 \cdot A^{11}B^{m-11}$	0	0	0	0	0	0	0	0	0	120
$E_4E_6D \cdot A^{11}B^{m-11}$	0	0	0	0	0	0	0	0	0	−672
$E_4^6 \cdot A^{12}B^{m-12}$	0	0	0	0	0	0	0	0	0	−11
$E_4^3D \cdot A^{12}B^{m-12}$	0	0	0	0	0	0	0	0	0	188
$D^2 \cdot A^{12}B^{m-12}$	0	0	0	0	0	0	0	0	0	−8

Table 2: Coefficients of $\psi^{\rm opt}_{0,m}$ in terms of the standard basis. Here $D := 2^{11}3^3\Delta$.

For **m=6** several new things happen, all connected with the fact that 6 is the first value of m having more than one prime factor. First of all, the formula (9.14), which worked for prime powers, now fails: $-2\mathcal{H}|\mathcal{V}_{2,6}^{(1)}$ is not a possible choice for $\Phi_{2,6}$. The reason is that, as we will see later in general, the form $\mathcal{H}|\mathcal{V}_{2,m}^{(1)}$ (for any m) is invariant under all Atkin-Lehner involutions, but $\Phi_{2,m}$ is not when m is not a prime power. The expression $-2\mathcal{H}|\mathcal{V}_{2,6}^{(1)}$ does, however, give the optimal choice of the eigencomponent $\Phi_{2,6}^{++}$ of $\Phi_{2,6}$ with eigenvalues $W_2 = W_3 = +1$. The other eigencomponent, corresponding to $W_2 = W_3 = -1$, is obtained by subtracting a multiple of the non-weak Jacobi form $\varphi_{2,6}^{-}$ given in (4.48) from $\pi_2^{-}(\Phi_{2,6}^{\text{opt}})$ (which is weakly holomorphic but not weak), and turns out to be simply a multiple of the mock Jacobi form $\mathcal{F}_6$ defined in §7.2 (Example 3), giving as our first choice of the "optimal" mock modular form

$$\Phi_{2,6}^{\text{I}} \;=\; -2\,\mathcal{H}\,|\,\mathcal{V}_{2,6}^{(1)} \;-\; \frac{1}{24}\,\mathcal{F}_6\;. \tag{9.16}$$

The second new phenomenon for $m = 6$ is that requiring the order of the poles to be as small as possible no longer fixes a unique optimal choice. The function $\Phi_{2,6}^{\text{I}}$ defined above has an even part that is strongly holomorphic (because $\mathcal{H}$ is) and an odd part that has discriminant $\Delta \geq -1$ (because $C^{(6)}(\Delta) = 0$ for $\Delta < -1$), but we can change this choice without disturbing the discriminant bound $\Delta \geq -1$ by adding to it an arbitrary multiple of the function

$$\mathcal{K}_6 = \frac{E_4AB^5 - 5E_6A^2B^4 + 10E_4^2A^3B^3 - 10E_4E_6A^4B^2 + (5E_4^3 - \frac{1}{4}D)A^5B - E_4^2E_6A^5}{12^5} \tag{9.17}$$

which is (up to a scalar factor, fixed here by requiring $\mathcal{K}_6 = y - 2 + y^{-1} + \mathrm{O}(q)$) the unique Jacobi form of weight 2 and index 6 satisfying the same discriminant bound.

Two particularly nice choices besides the function $\Phi_{2,6}^{\text{I}}$ are the functions

$$\Phi_{2,6}^{\text{II}} \;=\; \Phi_{2,6}^{\text{I}} \;-\; \frac{1}{24}\,\mathcal{K}_6 \qquad \Phi_{2,6}^{\text{III}} \;=\; \Phi_{2,6}^{\text{I}} \;+\; \frac{1}{4}\,\mathcal{K}_6 \tag{9.18}$$

The special properties of these two choices are that $\Phi_{2,6}^{\text{II}}$ is the unique choice with $c(\Phi; 0, 1) = 0$ and (hence) has the form $12^{-7}\sum_{i=0}^{7} f_{7,i}A^{i-1}B^{7-i}$ with $f_{7,i}$ satisfying (9.15), while $\Phi_{2,6}^{\text{III}}$ is the unique choice for which $c(n,r) = 0$ whenever $24n = r^2$. (This last choice is from some points of view the most canonical one.) Each of these three special choices, as we shall see below, has analogues for all m, although in general none of them is canonical, since for large m we can change them by adding cusp forms without affecting their defining properties and we do not know any criterion that selects a unique form.

The following table, which gives the first few Fourier coefficients $C(\Phi; \Delta,\, r \pmod{12})$ for the Jacobi form $\mathcal{K}_6$ and each of the five mock Jacobi forms $-2\mathcal{H}|\mathcal{V}_{2,6}^{(1)}$, $\mathcal{F}_6$, $\Phi_{2,6}^{\text{I}}$, $12\Phi_{2,6}^{\text{II}}$ and $\Phi_{2,6}^{\text{III}}$, summarizes the above discussion. In this table we have set in boldface the Fourier coefficients $C(\Delta, r)$ whose special properties determine the choice of Φ (namely, $C(-1,\pm1) + C(-1,\pm5) = 0$ for Φ^{I}, $C(-1,\pm1) = 0$ for Φ^{II}, and $C(0,0) = 0$ for Φ^{III}). We mention in passing

Δ	-1	-1	0	8	12	15	20	23	23	24	32	36	39
$r \pmod{12}$	±1	±5	0	±4	6	±3	±2	±1	±5	0	±4	6	±3
$\mathcal{K}_6$	1	1	-2	10	28	10	-12	13	13	-44	58	104	2
$-2\mathcal{H}\,\vert\,\mathcal{V}_{2,6}$	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1	-1	$-\frac{3}{2}$	$-\frac{3}{2}$	-1	$-\frac{5}{2}$	-2	-2
$\mathcal{F}_6$	-1	1	0	0	0	0	0	35	-35	0	0	0	0
$\Phi^{\rm I}_{2,6}$	$\boldsymbol{\frac{1}{24}}$	$-\boldsymbol{\frac{1}{24}}$	$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1	-1	$-\frac{71}{24}$	$-\frac{1}{24}$	-1	$-\frac{5}{2}$	-2	-2
$12\,\Phi^{\rm II}_{2,6}$	$\mathbf{0}$	-1	7	-11	-26	-17	-6	-42	-7	10	-59	-76	-25
$\Phi^{\rm III}_{2,6}$	$\frac{7}{24}$	$\frac{5}{24}$	$\mathbf{0}$	2	6	$\frac{3}{2}$	-4	$\frac{7}{24}$	$\frac{77}{24}$	-12	12	24	$-\frac{3}{2}$

Table 3: Jacobi and mock Jacobi forms of weight 2 and index 6

that the function $\psi^{\rm I}_{0,7} = 12^7 A\varphi^{\rm I}_{2,6}$ corresponding to (9.16) is precisely equal to B times $\psi^{\rm opt}_{0,6}$ as given in Table 2. We do not know if this has any special significance. It may be connected with the fact that $\psi^{\rm opt}_{0,6}$, which is equal to the quotient $\vartheta_1(\tau,4z)/\vartheta_1(\tau,2z)$ (Watson quintuple product identity), is a (-1)–eigenvector of the involution $W_2 = W_3$ on $\widetilde{J}_{0,6}$.

For **m=10** the situation is similar. Again, we can achieve $\Delta \geq -1$ but not uniquely because there is a weak Jacobi form $\mathcal{K}_{10} \in \widetilde{J}_{2,10}$ (unique up to a normalizing factor, which we fix in the same way as before) satisfying the same discriminant bound and we can change $\varphi_{2,10}$ an arbitrary multiple of $\mathcal{K}_{10}$. We again have three distinguished choices

$$\varphi^{\rm II}_{2,10} \;=\; \frac{\psi^{\rm II}_{0,11}}{12^{11}A}\,, \qquad \varphi^{\rm I}_{2,10} \;=\; \varphi^{\rm II}_{2,10} + \frac{1}{12}\mathcal{K}_{10} \qquad \varphi^{\rm III}_{2,10} \;=\; \varphi^{\rm I}_{2,10} + \frac{3}{8}\mathcal{K}_{10}\,, \tag{9.19}$$

where $\psi^{\rm II}_{0,11} \in \widetilde{J}_{0,11}$ is given (with $D = 32(E_4^3 - E_6^2) = 2^{11}3^3\Delta$ as in Table 2) by

$$\begin{aligned}\Psi^{\rm II}_{11} = {}& B^{11} - 55E_4A^2B^9 + 330E_6A^3B^8 - 990E_4^2A^4B^7 + 1848E_4E_6A^5B^6 - (2310E_4^3 - 111D)A^6B^5\\ &+1980E_4^2E_6A^7B^4 - (1155E_4^4 + 120E_4D)A^8B^3 + (440E_4^3E_6 + 20E_6D)A^9B^2\\ &-(99E_4^5 - 117E_4^2D)A^{10}B + (10E_4^4E_6 - 56E_4E_6D)A^{11}\end{aligned} \tag{9.20}$$

and is the unique choice satisfying (9.15), $\Phi^{\rm III}_{2,10}$ is the unique choice with no $\Delta = 0$ coefficients, and $\Phi^{\rm I}_{2,10}$ is the unique choice whose invariant part under W_2 (or W_5) is strongly holomorphic. Explicitly, we have

$$\Phi^{\rm I}_{2,10} \;=\; -\,2\,\mathcal{H}|\mathcal{V}^{(1)}_{2,10} - \frac{1}{12}\mathcal{F}_{10} \tag{9.21}$$

where $\mathcal{F}_{10}$ is a mock Jacobi form of weight 2 and index 10 with Fourier coefficients given by

$$c(\mathcal{F}_{10};n,r) = \begin{cases} \;\;C^{(10)}(40n - r^2) & r \equiv \pm1 \pmod{20} \text{ or } r \equiv \pm3 \pmod{20}\\ -\,C^{(10)}(40n - r^2) & r \equiv \pm7 \pmod{20} \text{ or } r \equiv \pm9 \pmod{20}\end{cases} \tag{9.22}$$

for a function $C^{10}(\Delta)$ whose first few values are

Δ	-1	31	39	71	79	111	119	151	159	191	199	231	239
$C^{(10)}(\Delta)$	-1	9	21	35	63	57	112	126	207	154	306	315	511

,

or equivalently, where $\mathcal{F}_{10}$ has a theta expansion of the form

$$\mathcal{F}_{10}(\tau,z) \;=\; \sum_{\substack{\ell \ (\mathrm{mod}\ 20)\\ (\ell,10)=1}} h_\ell(\tau)\,\vartheta_{10,\ell}(\tau,z) \tag{9.23}$$

with mock modular forms $h_\ell = h_\ell(\tau)$ having expansions beginning

$$h_1 \;=\; -h_9 \;=\; -h_{11} \;=\; h_{19} \;=\; q^{-1/40}\left(-1 \,+\, 21q \,+\, 63q^2 \,+\, 112q^3 \,+\, \cdots\right), \tag{9.24}$$

$$h_3 \;=\; -h_7 \;=\; -h_{13} \;=\; h_{17} \;=\; q^{31/40}\left(9 \,+\, 35q \,+\, 57q^2 \,+\, 126q^3 \cdots\right). \tag{9.25}$$

9.3 Further data and formulation of main results for the weight 2 family

In §9.2, we found that the functions $\Phi_{2,m}$ for $m \le 11$ could all be obtained, using the Hecke-like operators from §4.4, from just three functions $\mathcal{H}$, $\mathcal{F}_6$ and $\mathcal{F}_{10}$. Continuing, we find that for $\mathbf{m = 12}$ the same statement holds: the $(+1,+1)$ eigencomponent of this form with respect to the involutions W_4 and W_3 can be chosen to be $-2\mathcal{H}|\mathcal{V}^{(1)}_{2,12}$ and the $(-1,-1)$ eigencomponent can be chosen to be $\frac{1}{12}\mathcal{F}_6|\mathcal{V}^{(6)}_{2,2}$, with $\mathcal{V}^{(M)}_{k,t}$ as in (4.38) (so $V^{(1)}_{2,12} = \frac{1}{4}(V_{12} - 2V_2U_3)$ and $\mathcal{V}^{(6)}_{2,2} = \frac{1}{2}V_2$). Thus

$$\Phi_{2,12} \;=\; -2\mathcal{H}\,|\,\mathcal{V}^{(1)}_{2,12} \,+\, \frac{1}{12}\,\mathcal{F}_6\,|\,\mathcal{V}^{(6)}_{2,2} \tag{9.26}$$

is a possible choice for $\Phi_{2,12}$. Similarly, for $\mathbf{m = 18}$ and $\mathbf{m = 24}$ we can take the functions

$$\Phi_{2,18} \;=\; -2\mathcal{H}\,|\,\mathcal{V}^{(1)}_{2,18} \,+\, \frac{1}{24}\,\mathcal{F}_6\,|\,\mathcal{V}^{(6)}_{2,3}\,, \qquad \Phi_{2,24} \;=\; -2\mathcal{H}\,|\,\mathcal{V}^{(1)}_{2,24} \,+\, \frac{1}{24}\,\mathcal{F}_6\,|\,\mathcal{V}^{(6)}_{2,4} \tag{9.27}$$

as our representatives of the coset $\Phi^{\text{stand}}_{2,m} + \widetilde{J}_{2,m}$.

On the other hand, the indices $\mathbf{m = 14}$ and $\mathbf{m = 15}$ are similar to the indices $m = 6$ and $m = 10$ looked at previously: here we find optimal mock Jacobi forms of the form

$$\Phi_{2,14} \;=\; -2\mathcal{H}\,|\,\mathcal{V}^{(1)}_{2,14} \,-\, \frac{1}{8}\,\mathcal{F}_{14}\,, \qquad \Phi_{2,15} \;=\; -2\mathcal{H}\,|\,\mathcal{V}^{(1)}_{2,15} \,-\, \frac{1}{6}\,\mathcal{F}_{15}\,, \tag{9.28}$$

where $\mathcal{F}_{14}$ and $\mathcal{F}_{15}$ are forms anti-invariant under W_p $(p|m)$ with Fourier expansions of the form

$$c(\mathcal{F}_{14};n,r) = \begin{cases} \;\;C^{(14)}(56n - r^2) & \text{for } r \equiv \pm1,\ \pm3 \text{ or } \pm5 \pmod{28}, \\ -\,C^{(14)}(56n - r^2) & \text{for } r \equiv \pm9,\ \pm11, \text{ or } \pm13 \pmod{28}, \end{cases} \tag{9.29}$$

$$c(\mathcal{F}_{15};n,r) = \begin{cases} \;\;C^{(15)}(60n - r^2) & \text{for } r \equiv \pm1,\ \pm2,\ \pm4, \text{ or } \pm7 \pmod{30}, \\ -\,C^{(15)}(60n - r^2) & \text{for } r \equiv \pm8,\ \pm11,\ \pm13, \text{ or } \pm14 \pmod{30}, \end{cases} \tag{9.30}$$

with coefficients $C^{14}(\Delta)$ and $C^{15}(\Delta)$ given for small Δ by

Δ	-1	31	47	55	87	103	111	143	159	167	199	215	223
$C^{(14)}(\Delta)$	-1	3	10	15	15	30	42	20	51	65	51	85	120

,

Δ	-1	11	44	56	59	71	104	116	119	131	164	176	179
$C^{(15)}(\Delta)$	-1	2	10	4	14	9	28	24	29	26	40	28	54

.

This pattern continues: as we continue to higher values, we find that for each index M which is a product of an even number of distinct primes (i.e., for which $\mu(M)=1$, where $\mu(\,\cdot\,)$ denotes the Möbius function), we need to introduce a new function $\mathcal{Q}_M$ that is anti-invariant under W_p for all p dividing M, but that these functions and their images under Hecke-like operators then suffice to give possible choices for our special mock Jacobi forms in all indices, in the sense that the function

$$\Phi_{2,m} \;=\; \frac{1}{2^{\omega(m)-1}} \sum_{\substack{M|m\\ \mu(M)\,=\,+1}} \mathcal{Q}_M \,|\, \mathcal{V}^{(M)}_{2,\,m/M} \qquad (\,\omega(m) := \#\{p \mid p \text{ prime},\, p|M\}\,) \tag{9.31}$$

belongs to the coset $\Phi^{\text{stand}}_{k,m} + \widetilde{J}_{2,m}$ for all $m \geq 1$. (The factor $2^{1-\omega(m)}$ will be explained below.) In fact more is true. The functions $\mathcal{Q}_M$ in (9.31) are simply the $(-1,\dots,-1)$ eigencomponents of $\Phi_{2,M}$ for some choice of these latter functions within their defining cosets, and the various terms in (9.31) are the eigencomponents of $\Phi_{2,m}$ for the involutions W_{m_1} $(m_1\|m)$. In other words, we have the following theorem:

Theorem 9.1. *For any choice of the mock Jacobi forms $\Phi_{2,m} \in \Phi^{\text{stand}}_{2,m} + \widetilde{J}_{2,m}$ $(m \in \mathbb{N})$, the eigencomponents $\Phi^{\varepsilon_1,\dots,\varepsilon_r}_{2,m}$ $(m = \prod_i p_i^{\nu_i},\ \varepsilon_i = \pm 1,\ \prod \varepsilon_i = +1)$ with respect to the decomposition[19] (4.41) are given by*

$$\frac{1}{2}\Phi_{2,m}\,\Big|\,\prod_{i=1}^{r}\big(1+\varepsilon_i W_{p_i^{\nu_i}}\big) \;=\; \mathcal{Q}_M \,|\, \mathcal{V}^{(M)}_{2,m/M} \quad \big(\mathit{mod}\ (J^{!}_{2,m})^{\varepsilon_1,\dots,\varepsilon_r}\big)\,, \tag{9.32}$$

where M is the product of the p_i with $\varepsilon_i = -1$, $\mathcal{Q}_M$ is defined by

$$\mathcal{Q}_M \;=\; \frac{1}{2}\Phi_{2,M}\,\Big|\,\prod_{p|M}\big(1-W_p\big) \;=\; 2^{\omega(M)-1}\,\Phi^{-,\dots,-}_{2,M} \qquad (\mu(M)\;=\;+1), \tag{9.33}$$

and $\mathcal{V}^{(M)}_{2,m/M}$ is the Hecke-like operator defined in (4.38). In particular, the decomposition (9.31) is always true modulo $\widetilde{J}_{2,m}$, and if we choose the functions $\Phi_{2,M}$ and consequently $\mathcal{Q}_M$ arbitrarily for all M with $\mu(M)=1$, then the function defined by (9.31) gives an admissible choice of $\Phi_{2,m}$ for all indices $m \geq 1$.

---

[19] or rather, its analogue for elliptic forms

Examples. The functions $\mathcal{Q}_M$ for the first few values of M with $\mu(M) = 1$ are

$$\mathcal{Q}_1 = -\mathcal{H}\,, \quad \mathcal{Q}_6 = -\frac{\mathcal{F}_6}{12}\,, \quad \mathcal{Q}_{10} = -\frac{\mathcal{F}_{10}}{6}\,, \quad \mathcal{Q}_{14} = -\frac{\mathcal{F}_{14}}{4}\,, \quad \mathcal{Q}_{15} = -\frac{\mathcal{F}_{15}}{3}\,, \quad \dots\,, \tag{9.34}$$

where $\mathcal{H}, \mathcal{F}_6, \dots, \mathcal{F}_{15}$ are the functions introduced in §7.2, §9.2, and this subsection. (The reasons for the various normalizations, and for having two symbols $\mathcal{F}_M$ and $\mathcal{Q}_M$ to denote proportional functions, will be discussed later.)

We will prove Theorem 9.1 towards the end of this section by calculating the shadows of the mock Jacobi forms on the two sides of (9.32), since two mock Jacobi forms differ by a true Jacobi form if and only if they have the same shadow.

Remarks. 1. We wrote $\big(J^!_{2,m}\big)^{\varepsilon_1,\dots,\varepsilon_r}$ rather than $\widetilde{J}^{\varepsilon_1,\dots,\varepsilon_r}_{2,m}$ in the theorem because of the fact mentioned in §4.4 that the involutions W_{m_1} do not preserve the space of weak Jacobi forms (or weak elliptic forms), so that the left-hand side of eq. (9.32) for an arbitrary choice of $\Phi_{2,m}$ may contain monomials $q^n y^r$ with $n < 0$. In fact, it will follow from the results given below that the functions $\Phi_{2,M}$ can be chosen in such a way that the eigencomponents $\Phi^{-,\dots,-}_{2,M}$ do not contain any negative powers of q, in which case the expression on the right-hand side of (9.31) really does lie in $\Phi^{\text{stand}}_{k,m} + \widetilde{J}_{2,m}$ as asserted, and not merely in $\Phi^{\text{stand}}_{k,m} + J^!_{2,m}$.
2. An exactly similar statement is true for the functions $\Phi_{1,m}$, except that in that case we have $\prod_i \varepsilon_i = -1$ (because in odd weight W_m acts as multiplication by -1), so that the M that occur in the decomposition of $\Phi_{1,m}$ are the divisors of m with $\mu(M) = -1$, and the functions $\mathcal{Q}_M$ that occur in the decomposition of $\Phi_{1,M}$ are defined by

$$\mathcal{Q}_M \;=\; \frac{1}{2}\,\Phi_{1,M}\Big|\prod_{p|M}\big(1 - W_p\big) \;=\; 2^{\omega(M)-1}\,\Phi^{-,\dots,-}_{2,M} \qquad (\mu(M) \;=\; -1) \tag{9.35}$$

instead of equation (9.33). The reason for the factor $2^{\omega(m)-1}$ in these two equations is now apparent: it is the minimal factor making $\mathcal{Q}_M$ an integral linear combination of images of $\Phi_{k,m}$ under Atkin-Lehner involutions. (We could remove one factor 2 because for each decomposition $M = M_1M_2$, the two terms $\mu(M_1)\Phi_{k,m}|W_{M_1}$ and $\mu(M_2)\Phi_{k,m}|W_{M_2}$ are equal.)

Theorem 9.1 explains the first part of our observations, those connected with the eigencomponent decomposition of the mock Jacobi forms $\Phi_{2,m}$ and the fact that the collection of all of these functions can be expressed in terms of the special functions (9.34) and their images under Hecke-like operators. But it does not say anything about the orders of the poles of the functions in question, which was our original criterion for optimality. On the other hand, from the tables of Fourier coefficients of the functions $\mathcal{Q}_1, \dots, \mathcal{Q}_{15}$ that we have given we see that, although none of these functions except $\mathcal{Q}_1$ is strongly holomorphic, each of them satisfies the only slightly weaker discriminant bound $\Delta := 4mn - r^2 \geq -1$. This property will turn out to play a prominent role. We will call Jacobi or elliptic forms satisfying it **forms with optimal**

growth[20], because their Fourier coefficients grow more slowly than forms that do not obey this discriminant bound. We observe that any Jacobi or elliptic form having optimal growth is automatically a weak form, since $4nm - r^2 \geq -1$ implies $n \geq 0$.

As we continued our computations to higher values of M, we were surprised to find that this "optimal growth" property continued to hold for each M of the form pq (p and q distinct primes) up to 200. It followed that the full function $\Phi_{2,M}$ could also be chosen to have optimal growth in these cases, since $\Phi_{2,M} \;=\; \mathcal{Q}_1 | \mathcal{V}^{(1)}_{2,M} + \mathcal{Q}_M$ and the image of the strongly holomorphic form $\mathcal{Q}_1$ under $\mathcal{V}^{(1)}_{2,M}$ is strongly holomorphic. This made it particularly interesting to look at the case $\mathbf{m = 210}$, the first product of four distinct primes, because in this case, even if the function $\mathcal{Q}_{210} = \Phi^{----}_{2,210}$ can be chosen to be have optimal growth, then the six middle terms in the decomposition (9.31), which here takes the form

$$\begin{aligned} 8\,\Phi_{2,210} \;&=\; \mathcal{Q}_1 \,|\, \mathcal{V}^{(1)}_{2,210} \,+\, \mathcal{Q}_6 \,|\, \mathcal{V}^{(6)}_{2,35} \,+\, \mathcal{Q}_{10} \,|\, \mathcal{V}^{(10)}_{2,21} \,+\, \mathcal{Q}_{14} \,|\, \mathcal{V}^{(14)}_{2,15} \\ &\quad +\, \mathcal{Q}_{15} \,|\, \mathcal{V}^{(15)}_{2,14} \,+\, \mathcal{Q}_{21} \,|\, \mathcal{V}^{(21)}_{2,10} \,+\, \mathcal{Q}_{35} \,|\, \mathcal{V}^{(35)}_{2,6} \,+\, \mathcal{Q}_{210}\,, \end{aligned} \tag{9.36}$$

will not be. (Even if $\mathcal{Q}_M$ is chosen to have optimal growth, its image under the operator $\mathcal{V}^M_{2,t}$ will in general satisfy only $\Delta \geq -t^2$. For M of the form pq we had no problems with this because $\mathcal{Q}_1$, and hence also its image under $\mathcal{V}^{(1)}_{2,M}$, is actually holomorphic.) The corresponding computation was somewhat forbidding since the space $\widetilde{J}_{2,210}$ of forms by which we can change the initial choice $\Phi^{\text{stand}}_{2,210}$ has dimension 3815, and the conditions $C(\Delta, r) = 0$ for $\Delta < -1$ give 3813 linear constraints, so that we have to find the kernel of a roughly square matrix of size nearly 4000. Moreover, the coefficients of this matrix are gigantic, so that we have to work modulo some large prime. Nevertheless, this computation could be carried out (using PARI/GP), and it turned out that the function $\mathcal{Q}_{210}$ could again be chosen to have optimal growth! This solution is given by $\mathcal{Q}_{210} \;=\; -\mathcal{F}_{210}$, where $\mathcal{F}_{210} = \sum_{\ell \pmod{420}} h_\ell\, \vartheta_{210,\ell}$ with the Fourier coefficients of the mock theta functions h_ℓ given in Table 4. In this table, we have listed the values of ℓ with $(\ell, 210) = 1$ and $0 \leq \ell < 210$ (since $h_{420n \pm \ell} = h_\ell$ and $h_\ell = 0$ for $(\ell, 210) > 1$) in groups of 8, with $h_\ell = h_{\ell_0}$ for the first four and $h_\ell = -h_{\ell_0}$ for the second four, where ℓ_0 (shown in boldface) belongs to $\{1, 11, 13, 17, 19, 23\}$ and $h_{\ell_0} \;=\; \sum_{n \geq 0} C(n)\, q^{n - \ell_0^2/840}$. Thus, the third line of the table means that $h_{\pm 13} = h_{\pm 43} = \cdots = -h_{\pm 167} = -h_{\pm 197} \;=\; q^{-169/840}\big(6q + 17q^2 + 17q^3 + \ldots\big)$. A striking property of these coefficients is that they are so small: for instance, the largest value of $c(n, \ell)$ in the table is $c(12, 1) = 121$, with corresponding discriminant $\Delta = 4 \cdot 210 \cdot 12 - 1^2 = 10079$, while for the Mathieu form $h^{(2)}$ the table following eq. (7.31) gave the much larger coefficient 132825

---

[20] An alternative name would be "forms with only simple poles" because the forms $h_\ell(\tau) = \sum_\Delta C(\Delta, \ell)\, q^{\Delta/4m}$ have no singularity worse than $q^{-1/4m}$ as $q \to 0$, i.e., they have at most simple poles with respect to the local parameter $q^{1/4m}$ at $\tau = i\infty$. However, this terminology would be dangerous because it could also suggest simple poles with respect to z. Another reasonable name would be "nearly holomorphic," but this terminology also carries a risk of confusion since "almost holomorphic" is already a standard notation for modular forms which are polynomials in $1/\tau_2$ with holomorphic coefficients.

ℓ with $h_\ell = h_{\ell_0}$	ℓ with $h_\ell = -h_{\ell_0}$	$n=0$	1	2	3	4	5	6	7	8	9	10	11	12
1, 29, 41, 71	139, 169, 181, 209	-2	13	28	34	49	46	71	59	83	77	102	87	121
11, 31, 59, 101	109, 151, 179, 199	0	9	21	27	36	41	51	44	75	62	62	82	104
13, 43, 83, 97	113, 127, 167, 197	0	6	17	17	35	20	49	31	57	36	77	32	94
17, 53, 67, 73	137, 143, 157, 193	0	4	12	16	22	19	43	17	40	50	41	27	87
19, 61, 79, 89	121, 131, 149, 191	0	3	11	12	23	14	37	17	43	28	45	30	77
23, 37, 47, 103	107, 163, 173, 187	0	1	7	4	20	-1	32	3	30	10	50	-16	63

Table 4: Fourier coefficients of the function $\mathcal{F}_{210}$

for the much smaller discriminant $4 \cdot 2 \cdot 9 - 1^2 = 71$. This behavior is a direct consequence of the "optimal growth" property together with the fact that the index is large (cf. Theorem 9.3 below and the discussion in §9.5).

In fact, in this example even more happens: not only the function $\mathcal{Q}_{210} = \Phi_{2,210}^{----}$, but the whole function $\Phi_{2,M}$ (and hence each of its eigencomponents $\Phi_{2,210}^{\pm\pm\pm\pm}$) can be chosen to have optimal growth. It turns out that this statement holds for all square-free indices m, whether they have an even or an odd number of prime factors. If m is not square-free, then this is not true in general,[21] but it becomes true after a very small modification of the functions in question:

Theorem 9.2. *The functions* $\{\Phi_{2,m}\}_{m\geq 1}$ *can be chosen to have the form*

$$\Phi_{2,m}(\tau, z) \;=\; \sum_{d^2|m} \Phi^0_{2,m/d^2}(\tau,\, dz) \tag{9.37}$$

where $\{\Phi^0_{2,m}\}_{m\geq 1}$ *are primitive mock Jacobi forms having optimal growth.*

Corollary. *The functions* $\{\mathcal{Q}_M\}_{M>1,\,\mu(M)=1}$ *can be chosen to have optimal growth.*

(Here we have excluded the case $M = 1$ because the function $\mathcal{Q}_1 = -\mathcal{H}$ has already been chosen to have no poles at all.) Tables of Fourier coefficients of the functions $\mathcal{Q}_M$ appearing in the corollary for $M \leq 50$ are given in Appendix A.1.

If we denote by $\mathcal{E}_m^{\mathrm{OG}}$ the space of elliptic forms of index m (cf. §4.4) having optimal growth, then we have the sequence of inclusions

$$\mathcal{E}^0_m \;\subset\; \mathcal{E}_m \;\subset\; \mathcal{E}^{\mathrm{OG}}_m \;\subset\; \widetilde{\mathcal{E}}_m \;\subset\; \mathcal{E}^!_m \qquad \bigl(\varphi \in \mathcal{E}^{\mathrm{OG}}_m \underset{\mathrm{DEF}}{\iff} C_\varphi(\Delta, \ell) = 0 \text{ if } \Delta < -1\bigr)$$

of spaces of different types of elliptic forms (cusp, holomorphic, optimal growth, weak, weakly holomorphic) with different orders of growth. Theorem 9.2 says that we can represent all of our functions in terms of new, “primitive,” forms $\Phi^0_m \in \mathcal{E}^{\mathrm{OG}}_m$ for all m. The function Φ^0_m is unique up to the addition of an element of the space $J^{\mathrm{OG}}_{2,m} = \mathcal{E}^{\mathrm{OG}}_m \cap \widetilde{J}_{2,m}$ of Jacobi forms of weight 2 and index m of optimal growth. We will study this latter space in the next subsection.

---

[21] It is true if and only if m is the product of a square-free number and a prime power.

Theorem 9.2 has an immediate consequence for the growth of the Fourier coefficients of the optimal choice of $\Phi_{2,m}$, since the Hardy-Ramanujan circle method implies that the Fourier coefficients of a weakly holomorphic Jacobi (or mock Jacobi) form grow asymptotically like $e^{\pi\sqrt{\Delta|\Delta_0|}/m}$, where Δ_0 is the minimal value of the discriminants occurring in the form. We give a more precise version of this statement in the special case of forms of weight 2 having optimal growth.

Theorem 9.3. *Let φ be a weak or mock Jacobi form of weight 2 and index m with optimal growth. Then the Fourier coefficients of φ satisfy the asymptotic formula*

$$C_\varphi(\Delta, r) \;=\; \kappa_r \, e^{\pi\sqrt{\Delta}/m} \;+\; O\big(e^{\pi\sqrt{\Delta}/2m}\big) \tag{9.38}$$

as $\Delta \to \infty$, where the coefficient $\kappa_r = \kappa(\varphi; r)$ is given by

$$\kappa_r \;=\; -\sum_{\substack{\ell \ (\mathrm{mod}\ 2m) \\ \ell^2 \equiv 1 \ (\mathrm{mod}\ 4m)}} C_\varphi(-1, \ell)\, \cos\Big(\frac{\pi \ell r}{m}\Big)\,. \tag{9.39}$$

Proof. We only sketch the proof, since it is a standard application of the circle method. We write the coefficient $c_\varphi(\Delta, r)$ as a Fourier integral $\int_{-\frac{1}{2}+i\varepsilon}^{\frac{1}{2}+i\varepsilon} h_r(\tau)\, \mathbf{e}(-\Delta\tau/4m)\, d\tau$, with h_ℓ defined by (4.9) and (4.10). Its value is unchanged if we replace $h_r(\tau)$ by its completion $\widehat{h}_r(\tau)$, since they differ only by terms with negative discriminant The main contribution to the integral comes from a neighbourhood of $\tau = 0$, where we can estimate $\widehat{h}_r(\tau)$ to high accuracy as

$$\widehat{h}_r(\tau) \;=\; -\frac{(i/\tau)^{3/2}}{\sqrt{2m}} \sum_{\ell \ (\mathrm{mod}\ 2m)} \mathbf{e}\Big(\frac{r\ell}{2m}\Big)\, \widehat{h}_\ell\Big(-\frac{1}{\tau}\Big) \;=\; -\frac{\kappa_r}{\sqrt{2m}}\, (i/\tau)^{3/2}\, \mathbf{e}\Big(\frac{1}{4m\tau}\Big) \;+\; \mathrm{O}\big(\tau^{1/2}\big)\,,$$

the first equation holding by the S-transformation equation ([57], p. 59) and the second because the “optimal growth” assumption implies that $\widehat{h}_\ell(-1/\tau) = C_\varphi(-1, \ell)\, \mathbf{e}(1/4m\tau) + \mathrm{O}(1)$ as $\tau \to 0$ and the contribution from the Eichler integral of the shadow of $\widehat{h}_\ell$ is at most $\mathrm{O}(1/\tau)$. The rest follows by a standard calculation, the only point worthy of note being that since the h_ℓ's have weight 3/2 the Bessel functions that typically arise in the circle method are simple exponentials here. ☐

It is interesting to compare the statement of Theorem 9.3 with the tables given in the Appendix. This will be discussed in more detail in §9.5.

9.4 Optimal choice of the function $\Phi_{2,m}$

As we already saw, Theorem 9.2 fixes the choice of the mock Jacobi form $\Phi^0_{2,m}$ up to the addition of an element of the space $J^{\rm OG}_{2,m}$ of Jacobi forms of weight 2 and index m having optimal growth. The structure of this latter space is described by the following theorem, proved in §10.

Theorem 9.4. *The quotient $J^{\rm OG}_{2,m}/J_{2,m}$ has dimension 1 for every m. An element $\mathcal{K}_m \in J^{\rm OG}_{2,m}$ representing a non-trivial element of this quotient space can be chosen to satisfy*

$$\mathcal{K}_m(\tau,z) \;=\; y \,-\, 2 \,+\, y^{-1} \,+\, O(q) \quad \in\, \mathbb{C}[y,y^{-1}][[q]]\,,$$

and then has polar coefficients given by

$$C(\mathcal{K}_m;\,\Delta,\,r) \;=\; \begin{cases} 1 & \textit{if}\ \Delta=-1,\ r^2\equiv 1 \pmod{4m}\,, \\ 0 & \textit{if}\ \Delta<-1\,. \end{cases} \tag{9.40}$$

The function $\mathcal{K}_m$ can be chosen to be primitive and invariant under all the operators W_{m_1} $(m_1\|m)$.

The choice of $\mathcal{K}_m$ described in the theorem is unique up to the addition of a primitive holomorphic Jacobi form invariant under all W_m operators. This first occurs for $m=25$, so $\mathcal{K}_m$ is unique for small values of m. Because of Theorem 9.4, the function Φ^0_m has only a very small amount of choice: we can change it without disturbing the "optimal growth" property by adding an arbitrary multiple of $\mathcal{K}_m$ and an arbitrary holomorphic Jacobi form, but this is the only freedom we have. In particular, we have three special choices of "optimal" $\Phi^0_{2,m}$, each unique for $m<25$, corresponding to the three that we already encountered for $m=6$ and $m=10$ in §9.2:

(I) The $(+,\cdots,+)$-eigencomponent of $\Phi^{0,\rm I}_{2,m}$ is strongly holomorphic.

(II) $\Phi^{0,\rm II}_{2,m}$ has $c(0,1)=0$.

(III) $\Phi^{0,\rm III}_{2,m}$ has $c(0,0)=0$.

Each of these three choices has certain advantages over the other ones. In case (I) we can choose $\Phi^{0,\rm I}_{2,m}$ within its class modulo $J^0_{2,m}$ so that its $(+,\cdots,+)$-eigencomponent is precisely equal to $2^{1-\omega(m)}\mathcal{Q}_1|\mathcal{V}^{(1)}_{2,m}$ and hence has Fourier coefficients expressible in terms of class numbers, like in the examples for small m that we already saw. Also, $\Phi^{0,\rm I}_{2,m}$ is strongly holomorphic when m is a prime power, so that this is certainly the best of the three choices in that special case. The function occurring in (II) has a Fourier expansion beginning $\frac{1}{12}(m+1)+O(q)$, and the corresponding $\psi^{0,\rm II}_{m+1}$ satisfies (9.15) and belongs to the family of the weight zero, index m Jacobi forms discussed in [66]. The choice (III) looks at first sight like the least natural of the three, since it might seem strange to insist on killing a $\Delta = 0$ term when there are $\Delta = -1$ terms that are more singular, but in fact it is the nicest in some ways, since it can be further improved by the addition of a holomorphic Jacobi form to have *all* of its $\Delta = 0$ coefficients equal to zero, and its polar coefficients are then given by an elegant formula, which we now state.

Theorem 9.5. *The mock Jacobi form $\Phi^{0,\mathrm{III}}_{2,m}$ in Theorem 9.2 can be chosen, uniquely up to the addition of a Jacobi cusp form of weight 2 and index m, so that all its Fourier coefficients with $\Delta = 0$ vanish. Its polar coefficients are then given by*

$$C(\Phi^{0,\mathrm{III}}_{2,m}; -1, r) \;=\; \frac{1}{24}\sum_{d^2|m} \mu(d)\,\Bigl(\bigl(\frac{r-1}{2}, \frac{m}{d^2}\bigr) \,+\, \bigl(\frac{r+1}{2}, \frac{m}{d^2}\bigr)\Bigr) \quad \textit{for } r^2 \equiv 1 \pmod{4m}. \tag{9.41}$$

Moreover, the function $\Phi^{0,\mathrm{III}}_{2,m}$ can be chosen to be primitive without affecting these properties.

This also gives a multiplicative formula for the polar coefficients of all eigencomponents of $\Phi^{0,\mathrm{III}}_{2,m}$:

Corollary. *Write $m = \prod_{i=1}^{s} P_i$, where $P_i = p_i^{\nu_i}$ with distinct primes p_i and exponents $\nu_i \ge 1$. Then for any $\varepsilon_i \in \{\pm 1\}$ with $\varepsilon_1 \cdots \varepsilon_s = 1$, we have*

$$C\Bigl(\Phi^{0,\mathrm{III}}_{2,m}\,\Big|\,\prod_{i=1}^{s}\bigl(1 + \varepsilon_i W_{P_i}\bigr);\, -1,\, r\Bigr) \;=\; \frac{\pm 1}{12}\prod_{i=1}^{s}\begin{cases}(p_i + \varepsilon_i) & \textit{if } \nu_i = 1,\\ (p_i^{\nu_i} - p_i^{\nu_i - 2}) & \textit{if } \nu_i \ge 2,\end{cases} \tag{9.42}$$

where the sign is the product of the ε_i for all i for which $r \equiv -1 \pmod{2P_i}$. In particular, for the functions $\mathcal{Q}_M$ ($M > 1$, $\mu(M) = 1$), chosen by the Corollary to Theorem 9.2 to have optimal growth, the Fourier coefficients with non-positive discriminant are given by

$$C(\mathcal{Q}_M; \Delta, r) \;=\; \begin{cases} \quad 0 & \textit{if } \Delta = 0\,,\\ \pm\dfrac{\varphi(M)}{24} & \textit{if } \Delta \;=\; -1\,,\end{cases} \tag{9.43}$$

where in the second line $r^2 \equiv 1 \pmod{4M}$ and the sign is determined by $\mu((\frac{r+1}{2}, M)) = \pm 1$.

The final statement of the corollary can be compared with the examples in (9.34), in which each of the $\mathcal{F}_M$'s occurring (for $M > 1$) was normalized with $C(\mathcal{F}_M; -1, 1) = -1$, and the numerical factor relating $\mathcal{Q}_M$ and $\mathcal{F}_M$ equals $-\varphi(M)/24$ in each case. To give a feeling for the nature of these functions in general, and as an illustration of the formulas in Theorem 9.5 and its corollary, we give in Table 5 below the first few Fourier coefficients $C(\Phi; \Delta, \ell)$ for these and some related functions Φ for $m = 25$ (the first index where $J_{2,m} \neq \{0\}$ and hence $\mathcal{K}_m$ is not unique), $m = 37$ (the first index where $J^0_{2,m} \neq \{0\}$ and hence $\Phi^{0,\mathrm{III}}_{2,m}$ is not unique), $m = 50$ (the first case where (9.42) is non-trivial), and $m = 91$ (the first case where where $J^{0,--}_{2,m} \neq \{0\}$ and hence $\mathcal{Q}_m$ is not unique). The free coefficients "a", "b" and "c" in the table arise because

Δ	-1	0			4	11	16	19	24	31
$\pm\ell \pmod{50}$	1	0	10	20	14	17	22	9	24	13
$\mathcal{K}_{25}$	1	-2	a	$1-a$	$1-a$	$1+2a$	$2+3a$	$2-4a$	$5-2a$	$3+2a$
$\Phi^{0,\mathrm{III}}_{2,25}$	1	0	0	0	0	1	2	-1	2	1
$\Phi^{0,\mathrm{I}}_{2,25}$	0	2	c	$-1-c$	$-1-c$	$2c$	$3c$	$-3-4c$	$-3-2c$	$-2+2c$

Δ	-1	0	3	4	7	11	12	16	27	28
$\pm\ell \pmod{74}$	1	0	21	12	17	27	32	24	11	34
$\mathcal{K}_{37}$	1	-2	a	a	$1-a$	$1+a$	$2-a$	$2-2a$	$1-3a$	$2+3a$
$\Phi^{0,\mathrm{III}}_{2,37}$	$\frac{19}{12}$	0	b	$-\frac{1}{6}+b$	$\frac{1}{4}-b$	$\frac{11}{12}+b$	$\frac{3}{2}-b$	$1-2b$	$-\frac{3}{4}-3b$	$\frac{13}{6}+3b$
$\Phi^{0,\mathrm{I}}_{2,37}$	0	$\frac{19}{6}$	c	$-\frac{1}{6}+c$	$-\frac{4}{3}-c$	$-\frac{2}{3}+c$	$-\frac{5}{3}-c$	$-\frac{13}{6}-2c$	$-\frac{7}{3}-3c$	$-1+3c$

Δ	-1		0			4	16	24	31		36	39
$\pm\ell \pmod{100}$	1	49	0	20	40	14	28	24	13	37	42	19
$\mathcal{K}_{50}$	1	1	-2	a	$1-a$	$-a$	$1+a$	$1-2a$	$1+2a$	$1+2a$	$2+3a$	$-2a$
$\Phi^{0,\mathrm{III}}_{2,50}\lvert(1+W_2)$	3	3	0	0	0	-2	2	-2	3	3	8	-1
$\Phi^{0,\mathrm{III}}_{2,50}\lvert(1-W_2)$	1	-1	0	0	0	0	0	0	-1	1	0	-1

Δ	-1	3	12	27	40	48	55	68	75
$\pm\ell \pmod{182}$	$1\,/\,27$	$19\,/\,33$	$38\,/\,66$	$57\,/\,83$	$18\,/\,60$	$50\,/\,76$	$71\,/\,85$	$32\,/\,46$	$17\,/\,87$
$\mathcal{Q}_{91}$	3	a	$-1+a$	$-1-2a$	$-1-3a$	$-2+2a$	$3a$	$1+3a$	$-5+2a$

Table 5: Examples of non-unique Jacobi and mock Jacobi forms

$\dim J_{2,25} = \dim J^0_{2,37} = \dim J_{2,50} = \dim J^0_{2,91} = 1$. In the case of $m = 91$ we show the two ℓ-values corresponding to each value of Δ in the format $\ell\,/\,\ell^*$ and give in the following line the Fourier coefficient $C(\mathcal{Q}_{91};\Delta,\ell) = -C(\mathcal{Q}_{91};\Delta,\ell^*)$. In the cases of $m = 25$ and $m = 37$, we have given the (non-unique) holomorphic mock Jacobi form $\Phi^{0,\mathrm{I}}_{2,2m}$ as well as $\Phi^{0,\mathrm{III}}_{2,2m}$.

In summary, the "optimal growth" property of $\Phi^0_{2,m}$ and the near-uniqueness of the Jacobi form $\mathcal{K}_m$ have permitted us to pin down both of these functions in many cases, but there still remains an additive ambiguity whenever there are Jacobi cusp forms. We do not know how to resolve this ambiguity in general, and state this as an open problem:

Question: Is there a canonical choice of the forms $\mathcal{K}_m$ and $\mathcal{Q}_M$ for all positive integers m and all positive square-free integers M?

Should the answer to either of these questions be positive, the corresponding forms would be of considerable interest. In particular, one can speculate that the canonical choice of $\mathcal{Q}_M$, if it exists, might be canonically associated to the quaternion algebra of discriminant M, since there is a natural bijection between positive square-free integers and quaternion algebras over $\mathbb{Q}$, with the latter being indefinite or indefinite according as $\mu(M)=1$ (corresponding to our $k=2$ forms and to mock theta functions of weight 3/2) or $\mu(M)=-1$ (corresponding to $k=1$ and to mock theta functions of weight 1/2).

We end this subsection by describing one possible approach to answering the above question that we tried. This approach was not successful in the sense that, although it did produce a specific choice for $\mathcal{K}_m$ and $\Phi^0_{2,m}$ (or $\Phi^0_{1,m}$), these did not seem to have arithmetic coefficients and therefore had to be rejected, but the method of calculation involves a result that is of interest in its own right and may be useful in other contexts. The idea is to fix the choice of the Jacobi forms (or mock Jacobi forms) in question by demanding that they (or their completions) be orthogonal to Jacobi cusp forms with respect to the Petersson scalar product. (As an analogy, imagine that one wanted to determine the position of $E_{12}=1+\cdots$ in $M_{12}=\mathbb{C}\,E_4^3+\mathbb{C}\,\Delta$ but did not know how to compute the Fourier expansion of $\sum(m\tau+n)^{-12}$. Then one could simply write $E_{12}=E_4^3+a\Delta$ and use the orthogonality of Eisenstein series and cusp forms to compute the number $a=-432000/691$ numerically as $-(E_4^3,\Delta)/(\Delta,\Delta)$, where $(\;,\;)$ denotes the Petersson scalar product.) To apply this idea, we need a good way to compute the Petersson scalar product of Jacobi forms. This is the result of independent interest referred to above. For simplicity, we state it only in the case $k=2$.

Theorem 9.6. *Let φ_1 and φ_2 be two Jacobi cusp forms of weight 2 and index m, and set*

$$A(\Delta)\;=\sum_{\ell\;(\mathrm{mod}\;4m)} C_{\varphi_1}(\Delta,\ell)\,\overline{C_{\varphi_2}(\Delta,\ell)}\,,\qquad H(t)\;=\;\sum_{\Delta=1}^{\infty}\frac{A(\Delta)/\sqrt{\Delta}}{\exp(2\pi\sqrt{t\Delta/m})\;-\;1}\;. \tag{9.44}$$

Then the function $\dfrac{tH(t)-H(1/t)}{t-1}$ $(t\neq 1)$ has a constant value c_H which is proportional[22] *to the Petersson scalar product (φ_1,φ_2).*

Sketch of proof: We refer to [57] for the definition of (φ_1,φ_2) (eq. (13), p. 27) and for its expression as a simple multiple of $\sum_\ell(h_\ell^{(1)},h_\ell^{(2)})$, where $h_\ell^{(1)}$ and $h_\ell^{(2)}$ are the coefficients in the theta expansions of φ_1 and φ_2 (Theorem 5.3, p. 61). Calculating these weight 3/2 scalar products by the Rankin-Selberg method as explained in §3.1, we find that (φ_1,φ_2) can be expressed as a simple multiple of the residue at $s=3/2$ of the Dirichlet series $L(s)=\sum_{\Delta>0}A(\Delta)\Delta^{-s}$. The Rankin-Selberg method also shows that the function $\widetilde{L}(s)=2\,(2\pi)^{-s}m^s\,\Gamma(2s)\,\zeta(2s)\,L(s+\frac{1}{2})$ has a meromorphic continuation to all s, with poles only at $s=0$ and $s=1$, and is invariant under $s\to 1-s$. But $\widetilde{L}(s)$ is simply the Mellin transform of $H(t)$, so this implies the proposition by a standard argument. ☐

A rather surprising aspect of this theorem is that it is actually *easier* to compute Petersson scalar products for Jacobi forms than it is for ordinary modular forms of integral weight. In that case the analogue of $H(t)$ would have an infinite sum of K-Bessel functions multiplying each coefficient $A(\Delta)$, whereas here the Bessel functions are replaced by simple exponentials (because the modular forms $h_\ell^{(i)}$ have half-integral weight, so that the usual gamma factor $\Gamma(s)\Gamma(k+s-1)$ in the Rankin-Selberg method becomes a single gamma function $\Gamma(2s)$) and at the same time the infinite sum becomes a geometric series that can be summed explicitly.

---

[22] The exact constant of proportionality plays no role for our purposes, since we will only be concerned with ratios of Petersson scalar products. If the scalar product is normalized as in [57], its value is 1/2.

To apply the proposition to our situation in the case of $\mathcal{K}_{37}$, we start with an initial choice (like the one given in Table 5) and replace it by $\mathcal{K}_{37} + a\varphi_{37}$, where φ_{37} is the generator of $J^0_{2,37}$ (normalized by $C(\varphi_{37}\,;3,21) = 1$) and where $a \in \mathbb{R}$ is chosen to make this function orthogonal to φ_{37}, i.e., $a = -(\mathcal{K}_{37}, \varphi_{37})/(\varphi_{37}, \varphi_{37})$, in the hope that this a will be an integer or simple rational number. (Of course $\mathcal{K}_{37}$ is not a cusp form, but the integral defining the Petersson scalar product still converges,[23] the sum defining $H(t)$ converges for $t > \frac{1}{148}$ by Theorem 9.3, and Theorem 9.6 still holds.[24]) The numerical calculation for the case $\varphi_1 = \varphi_2 = \varphi_{37}$ gives the value $c_H = c_1 = -0.95284748097797917403$ (a value that can be checked independently, since up to a factor $-\pi$ it must be equal to the even period of the elliptic curve $y^2 - y = x^3 - x$, as calculated by GP-PARI as `ellinit`([0,0,1,-1,0])[15]). The corresponding number for $\varphi_1 = \varphi_{37}$ and $\varphi_2 = \mathcal{K}_{37}$ (with $\mathcal{K}_{37}$ chosen as in Table 5, with $a = 0$) is $c_2 = 0.26847600706319893417$, which is not a simple multiple of c_1, but when we calculate again with $\varphi_2 = \Phi^{0,\mathrm{III}}_{2,37}$ (again as in Table 5, with $b = 0$) we find a value $c_3 = 0.107471184190738587618$ that is related to the two previous values by $12c_3 - 19c_2 = 4c_1$ numerically to high precision. Thus in this case we have indeed found a form orthogonal to cusp forms and with rational coefficients. This at first looks promising, but in fact has a simple and disappointing explanation: the linear combination $\Phi^{0,\mathrm{III}}_{2,37} - \frac{19}{12}\mathcal{K}_{37} - \frac{1}{3}\varphi_{37} = \Phi^{0,\mathrm{I}}_{2,37} - \frac{1}{3}\varphi_{37}$ is nothing other than $\mathcal{Q}_1|V_{37}$ (as one can see in Table 5, where setting $c = -\frac{1}{3}$ gives entries equal to $-H(\Delta)$), and since the function $\mathcal{Q}_1 = -\mathcal{H}$ is constructed from an Eisenstein series, albeit a non-holomorphic one, it is automatically orthogonal to all cusp forms. And indeed, when we look at the case of $\mathcal{Q}_{91}$ and $\varphi_{91} \in J^{0,--}_{2,91}$, where there are no Eisenstein series in sight, then the calculation of the quotient $(\widehat{\mathcal{Q}}_{91}, \varphi_{91})/(\varphi_{91}, \varphi_{91})$ failed to yield a simple number. (The non-holomorphy of $\widehat{\mathcal{Q}}_{91}$ does not affect the application of Theorem 9.6, since its non-holomorphic part has Fourier coefficients only for $\Delta < 0$, where the coefficients of φ_{91} vanish.) As an incidental remark, we mention that the cusp form φ_{37}, whose first few

---

[23] The precise condition needed for the convergence of the Petersson scalar product of two weakly holomorphic Jacobi forms φ_1 and φ_2 is that $\mathrm{ord}_\infty(h_\ell^{(1)}) + \mathrm{ord}_\infty(h_\ell^{(2)}) > 0$ for all $\ell \in \mathbb{Z}/2m\mathbb{Z}$. This is amply satisfied in the case $\varphi_1 = \varphi_{37}$, $\varphi_2 = \mathcal{K}_{37}$, since the only negative value of $\mathrm{ord}_\infty(h_\ell^{(2)})$ is $-1/148$ for $\ell \equiv \pm 1 \pmod{74}$, and there $\mathrm{ord}_\infty(h_\ell^{(1)}) = 147/148$.

[24] The proof given above breaks down, since the numbers $A(\Delta)$ grow exponentially in $\sqrt{\Delta}$, so that the Dirichlet series $\sum_\Delta A(\Delta)/\Delta^s$ does not converge for any value of s, but the function $f(y) = y^{3/2}\sum_\Delta A(\Delta)e^{-\pi\Delta y/m}$, which is the constant term of an $SL(2,\mathbb{Z})$-invariant function that is small at infinity, still has a Mellin transform with a meromorphic continuation to all s, and if we define $\widetilde{L}(s)$ as the product of this Mellin transform with $\pi^{-s}\Gamma(s)\zeta(2s)$, then $\widetilde{L}(s)$ has the same analytic properties as before and its inverse Mellin transform $H(t)$ is still given by the sum in (9.44) for t sufficiently large and satisfies the same functional equation.

Fourier coefficients were obtained only with difficulty in [57] (pp. 118–120 and 145), can now be obtained very easily using the "theta blocks" of [69] as

$$\varphi_{37}(\tau,z) \;=\; \frac{1}{\eta(\tau)^6}\prod_{i=1}^{10}\vartheta(\tau,a_iz)\,, \qquad (a_1,\dots,a_{10}) \;=\; (1,1,1,2,2,2,3,3,4,5)\,,$$

and similarly $\varphi_{2,91}$ can be obtained easily as the difference of the two similarly defined theta blocks for the 10-tuples $(a_1,\dots,a_{10}) = (1,1,2,3,4,4,5,5,6,7)$ and $(1,1,2,2,3,3,4,5,7,8)$.

9.5 Observations on the weight one family, integrality, and positivity

In the last two subsections we discussed experimental properties of the family of weight 2 mock Jacobi forms that were introduced in §9.1 and formulated a number of general results that will be proved in §10. In this subsection we report on the experimental data for the weight 1 family and discuss various properties of both families that play a role in connection with other investigations in mathematics and physics (like "Umbral Moonshine") or that seem to be of independent interest.

The discussion of the weight 1 family is quite different in nature from that of the weight 2 case, because here there do not seem to be general structural results as there were there, but only a small number of special functions for small values of the index having nice properties. In particular, the analogue of Theorem 9.2 no longer holds, and we find instead that (at least up to $m = 100$, and presumably for all m) the appropriately defined primitive mock Jacobi forms $\Phi^0_{1,m}$ can be chosen to have optimal growth only for the indices

$$m \;=\; 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 12,\ 13,\ 16,\ 18,\ 25\,. \tag{9.45}$$

Similarly, although the analogue of Theorem 9.1 is still valid in the weight 1 case (now with the special mock Jacobi forms $\mathcal{Q}_M$ labelled by square-free integers M having an odd rather than an even number of prime factors), the functions $\mathcal{Q}_M$ can apparently be chosen to have optimal growth only for

$$M \;=\; 2,\ 3,\ 5,\ 7,\ 13,\ 30,\ 42,\ 70,\ 78\,. \tag{9.46}$$

(This list has a precise mathematical interpretation; see *3.* below.) Thus here a small number of special examples is singled out. Interestingly enough, these turn out to include essentially all of the special mock theta functions that have been studied in the literature in the past.

For convenience, we divide up our discussion into several topics.

1. Optimal choices for the weight one family

As in the weight 2 case, we begin our discussion by looking at small values of m. We will consider the original forms $\Phi_{1,m}$ first, starting with the "standard" choices (9.2) and trying to modify them by the addition of weak Jacobi forms to attain the property of having optimal growth. Actually, it is better to look at the primitive mock Jacobi forms, now defined by

$$\Phi^0_{1,m} \;\equiv\; \sum_{d^2|m} \frac{\mu(d)}{d}\, \Phi_{1,m/d^2}|U_d\,, \qquad \Phi_{1,m} \;\equiv\; \sum_{d^2|m} \frac{1}{d}\, \Phi^0_{1,m/d^2}|U_d \qquad (\mathrm{mod}\ \widetilde{J}_{1,m})\,, \tag{9.47}$$

instead of (9.37) and (10.52), for the same reason for which we introduced the functions $\Phi^0_{2,m}$ in §9.3: residue arguments similar to those that we will give in §10 for the weight 2 case show here that $\Phi_{1,m}$ can never have optimal growth unless it agrees with $\Phi^0_{1,m}$ (which happens if and only if m is square-free or the square of a prime number), so that the latter class is certain to be more productive. Finally, we will look at the forms $\mathcal{Q}_M$ where M is a product of an odd number of distinct prime factors, since according to Theorem 9.1 all of the forms $\Phi_{1,m}$ can be constructed from these. Unlike the situation for $k = 2$, here we do not have to worry about any choices, because of the following theorem (to be compared with Lemma 2.3 of [25] and to be proved, like all its companions, in §10) and corollary. The only issue is therefore whether the "optimal growth" condition can be fulfilled at all.

Theorem 9.7. *There are no weight one weak Jacobi forms of optimal growth.*

Corollary. *If any of the mock Jacobi forms $\Phi_{1,m}$, $\Phi^0_{1,m}$, or $\mathcal{Q}_M$ ($\mu(M) = -1$) can be chosen to have optimal growth, then that choice is unique.*

We now look at some small individual indices. For **m=1** there is nothing to discuss, since the function $\Phi^{\mathrm{st}}_{1,1} = (C/2A)^{\mathrm{F}}$ vanishes identically, as already explained in §8.5 (Example 1). (This, by the way, is the reason why $\Phi_{1,m} = \Phi^0_{1,m}$ when m is the square of a prime.) The case of **m=2** has also already been treated in the same place (Example 2), where we saw that

$$\varphi^{\mathrm{opt}}_{1,2} \;=\; \frac{B}{24\,A}\,C\,, \qquad \Phi^{\mathrm{opt}}_{1,2} \;=\; -\frac{1}{24}\,\mathcal{F}_2\,, \tag{9.48}$$

with $\mathcal{F}_2$ the mock modular form defined in §7 and related to the Mathieu moonshine story. For **m=3**, we find that the (unique) choice of $\Phi_{1,m}$ having optimal growth is given by

$$\varphi^{\mathrm{opt}}_{1,3} \;=\; \frac{B^2 - E_4A^2}{288A}\,C\,, \qquad \Phi^{\mathrm{opt}}_{1,3} \;=\; \varphi^{\mathrm{opt}}_{1,3} \;-\; \mathcal{A}_{1,3} \;=\; \frac{1}{12}\,\mathcal{F}_3\,, \tag{9.49}$$

where the normalizing factor 1/12 has been chosen to give $\mathcal{F}_3$ integral Fourier coefficients. These coefficients are determined by virtue of their periodicity and oddness by the values $C^{(3)}(\Delta) = C(\mathcal{F}_3; \Delta, r)$ for $r \in \{1, 2\}$, the first few values of $C^{(3)}(\Delta)$ being given by

Δ	-1	8	11	20	23	32	35	44	47	56	59
$C^{(3)}(\Delta)$	-1	10	16	44	55	110	144	280	330	572	704

.

(A more extensive table of these coefficients, listed separately for the two values $r = 1$ and $r = 2$, is given in Appendix A.2, where we also give the corresponding values for the other forms treated here, so that we no longer will give tables of coefficients in the text.) For **m=4** and **m=5**, the functions $\Phi_{1,m}$ can again be chosen (of course uniquely, by the corollary above) to have optimal growth, the corresponding meromorphic Jacobi forms $\varphi^{\text{opt}}_{1,4}$ and $\varphi^{\text{opt}}_{1,5}$ being given by

$$\frac{B^3 - 3E_4A^2B + 2E_6A^3}{2\cdot 12^2\,A}\,C \quad \text{and} \quad \frac{B^4 - 6E_4A^2B^2 + 8E_6A^3B - 3E_4^2A^4}{2\cdot 12^3\,A}\,C\,, \tag{9.50}$$

respectively. Note that in the case of $m = 4$, according to the decomposition result given in Theorem 9.1, we could also make the choice $\mathcal{Q}_2|\mathcal{V}^{(2)}_{1,2}$ for the function $\Phi_{1,4}$, but then it would have the worse discriminant bound $\Delta_{\min} = -4$. We also observe that, although 4 is not square-free, we do not have to distinguish between $\Phi_{1,m}$ and $\Phi^0_{1,m}$ here, because they differ by the function $\Phi^0_{1,1}|U_2$ and the function $\Phi_{1,1}$ vanishes identically.

It is now very striking that the numerators of the fractions in (9.49) and (9.50) are the *same* as the expressions that we already saw in eqs. (9.10), (9.11) and (9.12) in §9.2, meaning that for these values of m we have

$$\varphi^{\text{opt}}_{1,m} \;=\; \frac{C}{A}\,\psi^{\text{opt}}_{0,m-1} \;=\; C\,\varphi^{\text{opt}}_{2,m-2} \tag{9.51}$$

with the same weight 0 weak Jacobi modular forms $\psi^{\text{opt}}_{0,m-1}$ as were tabulated in Table 2. The same thing happens for several other small values of the index. We do not know the deeper reason (if any) for this coincidence, but merely remark that the weight 0 weak Jacobi forms obtained from the weight 1 meromorphic Jacobi forms in our family by dividing by C also play a prominent role in the recent work of Cheng, Duncan and Harvey [25] on "umbral moonshine," and also that the the forms $\psi^{\text{opt}}_{0,m}$ occur in a somewhat different context in earlier work of Gritsenko[66] related to Borcherds products, as was already mentioned in §9.3 when we introduced the "second standard choice" $\Phi^{0,\text{II}}_{2,m}$ of the primitive mock Jacobi form $\Phi^0_{2,m}$.

Continuing to larger values, we look first at primes, since for m prime the three functions of interest $\Phi_{1,m}$, $\Phi^0_{1,m}$ and $\mathcal{Q}_m$ all coincide. For **m=7** we find that $\mathcal{Q}_7 = \Phi^{\text{opt}}_{1,7}$ again has optimal growth and again satisfies (9.51) with $\varphi^{\text{opt}}_{0,6}$ as given in Table 2. However, for **m=11** the function $\mathcal{Q}_M$ *cannot* be chosen to be of optimal growth; the best possible choice, shown in the table A.3 in the appendix, has minimal discriminant -4. For **m=13** the function $\mathcal{Q}_{13}$ again has optimal growth and is again given by (9.51), but for larger primes this never seems to happen again, the minimal value of Δ for the optimal choices of $\mathcal{Q}_M$ for the first few prime indices being given by

M	2	3	5	7	11	13	17	19	23	29	31	37
$\Delta_{\min}$	1	1	1	1	4	1	4	4	8	5	8	4

The next simplest case is $m = p^2$, since here $\Phi_{1,m}$ is still primitive. For **m=9** and **m=25** this function can be chosen to have optimal growth, but apparently not after that. The following case to look at is $m = p^\nu$ with $\nu \geq 2$, since here the eigendecomposition (9.31) of $\Phi_{1,m}$ is trivial. (The eigenvalue of W_m must be -1 because the weight is odd, so $\Phi_{1,m} = \Phi^-_{1,m} = \mathcal{Q}_p|\mathcal{V}^{(p)}_{p^{\nu-1}}$.) Here we find optimal growth only for **m=8** and **m=16**. Next, for M square-free but not prime, where the decomposition (9.2) is trivial, the only two cases giving optimal growth seem to be **m=6** and **m=10**. And finally, for M neither square-free nor a power of a prime we also find two examples **m=12** and **m=18**. Altogether this leads to the list (9.45) given above. The first few coefficients of the forms $\Phi^0_{1,m}$ for all of these values of m are tabulated in Appendix A.2.

However, this is not the end of the story. If $\Phi_{1,m}$ has optimal growth for some m, then of course all of its eigencomponents with respect to the Atkin-Lehner operators also do, but the converse is not true, so there still can be further square-free values of M for which the function $\mathcal{Q}_M = 2^{1-\omega(M)}\,\Phi^{-,\ldots,-}_{1,M}$ has optimal growth, and this indeed happens. Since primes have already been treated and products of five or more primes are beyond the range of our ability to compute, we should consider products of three primes. Here we find four examples, for the indices

$$30 \;=\; 2\cdot 3\cdot 5\,, \qquad 42 \;=\; 2\cdot 3\cdot 7\,, \qquad 70 \;=\; 2\cdot 5\cdot 7\,, \qquad 78 \;=\; 2\cdot 3\cdot 13\,,$$

but for no other values less than 200 (and probably for no other values at all), giving the list (9.46). Again the Fourier coefficients of the corresponding forms, as well as those for all other $M < 150$ of the form $p_1p_2p_3$, have been listed in the appendix (part A.3). These examples are of particular interest, for a reason to which we now turn.

2. Relationship to classical mock theta functions

Recall from §7.1 that a mock modular form is called a mock theta function if its shadow is a unary theta series, and that this definition really does include all of Ramanujan's original mock theta functions that began the whole story. The theta expansion coefficients $h_\ell(\tau)$ of our special mock Jacobi forms $\Phi_{1,m}$ and $\Phi_{2,m}$ are always mock theta functions, since their shadows are multiples of the unary theta series $\varphi^1_{m,\ell}$ and $\varphi^0_{m,\ell}$, respectively. In fact, it turns out that many of our examples actually coincide with the mock theta functions studied by Ramanujan and his successors. This is the "particular interest" just mentioned.

Let us consider first the case **m=30**, the first of our $\mathcal{Q}_M$ examples with composite M. Because it is so highly composite, there are very few orbits of the group of Atkin-Lehner involutions on the group of residue classes modulo $2m$ and hence very few components in the theta decomposition of $\mathcal{Q}_{30}$. In fact there are only two, so that up to sign we get only two distinct coefficients h_1 and h_7 and $\mathcal{Q}_{30}$ has a decomposition

$$\begin{aligned} -3\,\mathcal{Q}_{30} \;=\;& h_1\left(\vartheta_{30,1}+\vartheta_{30,11}+\vartheta_{30,19}+\vartheta_{30,29}-\vartheta_{30,31}-\vartheta_{30,41}-\vartheta_{30,49}-\vartheta_{30,59}\right)\\ &+\; h_7\left(\vartheta_{30,7}+\vartheta_{30,13}+\vartheta_{30,17}+\vartheta_{30,23}-\vartheta_{30,37}-\vartheta_{30,43}-\vartheta_{30,47}-\vartheta_{30,53}\right). \end{aligned}$$

(The normalizing factor -3 is included for convenience: the function $-3\mathcal{Q}_{30}$ is the function $\mathcal{F}_{30}$ defined in 9.52 below.) The mock theta functions h_1 and h_7 have Fourier expansions beginning

$$\begin{aligned} h_1(\tau) &= q^{-1/120}\,(-1 + q + q^2 + 2q^3 + q^4 + 3q^5 + \cdots)\,, \\ h_7(\tau) &= q^{71/120}\,(1 + 2q + 2q^2 + 3q^3 + 3q^4 + 4q^5 + \cdots)\,. \end{aligned}$$

These expansions are easily recognized: up to trivial modifications they are two of the "mock theta functions of order 5" in the original letter of Ramanujan to Hardy, namely

$$\begin{aligned} 2 + q^{1/120}\,h_1(\tau) &= \chi_1(q) = \sum_{n=0}^{\infty} \frac{q^n}{(1-q^{n+1})\cdots(1-q^{2n})}\,, \\ q^{-71/120}\,h_2(\tau) &= \chi_2(q) = \sum_{n=0}^{\infty} \frac{q^n}{(1-q^{n+1})\cdots(1-q^{2n+1})}\,. \end{aligned}$$

(Actually, Ramanujan gave five pairs of mock theta functions of "order 5" in his letter, any two differing from one another by weakly holomorphic modular forms, but it was pointed out in [129] that of the five associated pairs of mock modular forms only $(q^{-1/120}(2-\chi_1(q)),\, q^{71/120}\chi_2(q))$ has the property that its completion transforms as a vector-valued modular form for the full modular group.) For **m=42**, exactly the same thing happens, and the three distinct components h_1, h_5 and h_{11} of $\mathcal{Q}_{42}$, whose initial coefficients are tabulated in Appendix A.3, turn out to be, up to powers of q and signs, identical with Ramanujan's three "order 7" mock theta functions $\mathcal{F}_{7,i}$ ($i = 1, 2, 3$) that were given as our first examples of mock modular forms in §7.1. What's more, this behavior continues: in [129] it was mentioned that one could construct mock theta functions of arbitrary prime[25] "order" p generalizing these two examples as quotients by $\eta(\tau)^3$ of certain weight 2 indefinite theta series, and an explicit example

$$M_{11,j}(\tau) = \frac{1}{\eta(\tau)^3} \sum_{\substack{m>2|n|/11 \\ n\equiv j \ (\mathrm{mod}\ 11)}} \Big(\frac{-4}{m}\Big)\Big(\frac{12}{n}\Big)\big(m\,\mathrm{sgn}(n) - \frac{n}{6}\big)\, q^{m^2/8-n^2/264} \qquad (j \in \mathbb{Z}/11\mathbb{Z})$$

was given for $p = 11$. Computing the beginnings of the Fourier expansions of these five mock theta functions (only five rather than eleven because $M_{11,j} = -M_{11,-j}$) and comparing with the table in Appendix A.3, we find perfect agreement with the coefficients in the theta expansion of $\mathcal{Q}_{66}$, the first example in that table of a form that is not of optimal growth.

This discussion thus clarifies the precise mathematical meaning of the number designated by Ramanujan as the "order" of his mock theta functions, but which he never defined: it is indeed related to the *level* of the completions of the corresponding mock modular forms, as was clear from the results of Zwegers and other authors who have studied these functions, but it is even more closely related to the *index* of the mock Jacobi forms that have these mock theta functions as the coefficients of its theta expansion.

---

[25] actually, only $(p, 6) = 1$ is needed

Of course the above cases are not the only ones where there is a relationship between the special mock Jacobi forms $\mathcal{Q}_M$ and classically treated mock theta functions, but merely the ones where the relationship between the "order" and the index is clearest. Without giving any further details, we say only that Ramanujan's order 3 functions are related in a similar way[26] to our $\mathcal{Q}_3$ and that his order 10 functions (not given in his letter to Hardy, but contained in the so-called "Lost Notebook") are related to $\mathcal{Q}_5$. We also refer once again to the beautiful recent work of Cheng, Duncan and Harvey [25] on "umbral moonshine," in which many of these same functions are discussed in detail in connection with representations of special finite groups.

Finally, we should mention that all of Ramanujan's functions were q-hypergeometric series but that in general no q-hypergeometric expression for the theta expansion coefficients of the functions $\Phi_{k,m}$ or $\mathcal{Q}_M$ is known. This seems to be an interesting subject for further research. One could look in particular for representations of this sort for the functions $\mathcal{Q}_{13}$ and $\Phi^0_{1,25}$ as given in Table A.2, since these have particularly small and smoothly growing coefficients that suggest that such a representation, if it exists, might not be too complicated. It would also be reasonable to look at $\mathcal{Q}_{70}$ and $\mathcal{Q}_{78}$, the only two other known weight 1 examples with optimal growth.

3. Integrality of the Fourier coefficients

In Subsections 9.2 and 9.3 and above, we encountered several examples of functions $\mathcal{Q}_M$ of optimal growth that after multiplication by a suitable factor gave a function $\mathcal{F}_M$ having all or almost all of the following nice arithmetic properties:

(A) All Fourier coefficients $C(\mathcal{F}_M;\Delta,r)$ are integral.

(B) The polar Fourier coefficient $C(\mathcal{F}_M;-1,1)$ equals -1.

(C) The non-polar Fourier coefficients $C(\mathcal{F}_M;\Delta,r)$ have a sign depending only on r.

(D) $C(\mathcal{F}_M;\Delta,r)$ is always positive for $r=r_{\min}$, the smallest positive number whose square equals $-\Delta \pmod{4m}$.

In particular, for $k=2$ this occurred for $M=6$, 10, 14 and 15 (compare eqs. (7.28), (9.22), (9.29) and (9.30), which describe the coefficients $C^{(M)}(\Delta)=C(\mathcal{F}_M;\Delta,r_{\min})$ for these indices) and for $k=1$, for the values $M=2$, 3, 5, 7, 13 and 30 (with "non-negative" instead of "positive" in (C) in the case of $M=13$), as well as for the functions $\Phi^0_{1,m}$ for the prime powers $M=4$, 8, 9, 16 and 25 (again with "positive" replaced by "non-positive," and also allowing half-integral values for the Fourier coefficients with $\Delta=0$). It is natural to ask for what values of M these properties can be expected to hold. We will discuss properties (A) and (B) here and properties (C) and (D) in point *4.* below.

---

[26] but not "on the nose," for instance, our function $h_{3,1}$ differs from Ramanujan's 3rd order mock theta function $-3\,q^{1/12}f(q^2)$ by an eta product.

Clearly either (A) or (B) can always be achieved by multiplying $\mathcal{Q}_M$ by a suitable scalar factor. We define a renormalized function $\mathcal{F}_M$ in all cases by setting

$$\mathcal{F}_M(\tau, z) \;=\; -c_M\, \mathcal{Q}_M(\tau, z)\,, \tag{9.52}$$

where c_M is the unique positive rational number such that $\mathcal{F}_M$ has integer coefficients with no common factor. Then (A) is always true and (B) and (C) will hold in the optimal cases. This normalization is also convenient for the presentation of the coefficients in tables, since there are no denominators and very few minus signs. Tables of the first few Fourier coefficients of all $\mathcal{F}_M$ up to $M = 50$ in the weight 2 case (i.e., for M with $\mu(M) = 1$) and of all known $\mathcal{F}_M$ having optimal growth in the weight 1 case (i.e., for M with $\mu(M) = -1$) are given in A.1 and A.3 of the Appendix, respectively. Looking at these tables, we see that the values of M for which all three properties above hold are $M = 6$, 10, 14, 15, 21 and 26 in the case $k = 2$ and $M = 2$, 3, 5, 7, 13, 30 and 42 in the case $k = 2$, and that in the $k = 2$ case we also get $\mathcal{F}_M$ satisfying (C) but not (B) for $M = 22$ and $\mathcal{F}_M$ satisfying (B) but not (C) for $M = 35$ and 39. This list is not changed if we look at a larger range of M (we have calculated up to $M = 210$ in the weight 2 case and $M = 100$ in the weight 1 case) or Δ (in each case we have calculated about twice as many Fourier coefficients as those shown in the tables), so that it seems quite likely that it is complete.

For the weight 2 functions, the corollary of Theorem 9.41 tells us that $C(\mathcal{F}_M; -1, 1)$ will always be negative with the normalization (9.52) and that we have

$$c_M \;=\; t \cdot \mathrm{denom}\Big(\frac{\varphi(M)}{24}\Big)\,, \qquad C(\mathcal{F}_M; -1, 1) \;=\; -t \cdot \mathrm{numer}\Big(\frac{\varphi(M)}{24}\Big) \tag{9.53}$$

for some positive rational number t. In all cases we have looked at, t turned out to be an integer, and in fact usually had the value 1. For instance, the second statement is true for all values of M tabulated in Appendix A, and also for $M = 210$, where Table 4 in §9.3 shows that $c_M = 1$ and hence $t = 1$, but the table below giving all values up to $M = 100$ shows that it fails for $M = 65$ and $M = 85$.

M	6	10	14	15	21	22	26	33	34	35	38	39	46	51	55	57	58	62	65	69	74	77	82	85	86	87	91	93	94	95
c_M	12	6	4	3	2	12	2	6	3	1	4	1	12	3	3	2	6	4	2	6	2	2	3	6	4	3	1	2	12	1
$C(\mathcal{F}_M; -1, 1)$	1	1	1	1	1	5	1	5	2	1	3	1	11	4	5	3	7	5	4	11	3	5	5	16	7	7	3	5	23	3
t	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	1	1	1	1

If it is true that t is integral, then equation (9.53) shows that the above list of $\mathcal{F}_M$ satisfying property (B) is complete, because we can then only have $|C(\mathcal{F}_M; -1, 1)| = 1$ if $\varphi(M)$ divides 24, which in turn implies that M is at most 90 and hence within the range of our calculations. We do not know whether the integrality is true in general, but in any case $12^{M+1}t$ must be an integer, since correcting the initial choice $\Phi_{2,M}^{\text{stand}} = 12^{-M-1}B^{M+1}A^{-1} - \mathcal{A}_{2,M}$ by weak Jacobi forms cannot make its denominator any worse.

Exactly similar comments apply in the weight 1 case. Here a residue calculation similar to the ones described in the next section (but with the residue $\mathcal{R}$ as defined in (10.17) below replaced by its weight 1 analogue as defined in (10.59)) tells us that in the case of optimal growth eq. (9.43) still holds, so that we find that eq. (9.53) still holds for some postive rational number t. If this t turns out to be an integer, it again follows that property (B) can only hold if $\varphi(M)|24$, and the list of all integers M with $\mu(M) = -1$ satisfying this latter condition coincides exactly with the list (9.46). Thus, although we cannot yet prove that that list is complete, we can at least understand where it comes from.

4. Positivity of the Fourier coefficients

We now turn to properties (C) and (D). Here we can use the asymptotic formula given in Theorem 9.3 to understand the question of the signs and asymptotics of the Fourier coefficients. (We will discuss only the weight 2 case in detail, but a similar theorem and similar discussion could equally well be given for the case $k = 1$.) This theorem implies that (C) "almost holds" in the sense that for any fixed value of r the Fourier coefficients $C(\Delta, r)$ eventually have a constant sign, so that there are at most finitely many exceptions to the statement. We therefore have to consider (D) and also the question of which values of r are most likely to give rise to the exceptional signs.

For the function $\mathcal{Q}_M$ (which we remind the reader differs from $\mathcal{F}_M$ by a negative proportionality factor, so that the expected behavior (D) would now say that $C(\Delta, r)$ is always negative when $r = r_{\min}$ is the smallest positive square root of $-\Delta$), the value of the constant κ_r appearing in that theorem as given according to (9.39) and (9.43) by

$$\kappa(\mathcal{Q}_M, r) \;=\; -\frac{\varphi(M)}{24} \cdot 2 \sum_{\substack{0<\ell<M \\ \ell^2 \equiv 1 \pmod{4M}}} \mu\Big(\big(\frac{\ell+1}{2}, M\big)\Big) \cos\Big(\frac{\pi \ell r}{M}\Big)\,. \tag{9.54}$$

A similar formula applies in the weight 1 case with "cos" resplaced by "sin." In the case when M is a prime, this immediately gives the positivity of $\kappa(\mathcal{Q}_M, r_{\min})$, and in the case when M has three prime factors, the tables in A.3 show that (D) usually fails. We therefore consider the weight 2 case, where $\mu(M) = +1$.

If M (which we always assume to be larger than 1) is a product of two primes, and in particular for all $M < 330$ with $\mu(M) = 1$ except 210, the sum on the right-hand side has only two terms and reduces to simply $\cos(\pi r/M) - \cos(\pi r^*/M)$, and since $0 < r_{\min} < r^*_{\min} < 2M$ and $\cos(\pi x)$ is decreasing on the interval $[0, \pi]$, the value of $-\kappa(\mathcal{Q}_M, r_{\min})$ is indeed always strictly negative. It then follows from Theorem 9.3 that almost all $C(\mathcal{F}_M; \Delta, r_{\min})$ are positive, but there can be, and in general are, finitely many exceptions. However, for small M these are fairly rare, e.g., the following table lists all triples $(M,\ r = r_{\min},\ n)$ within the range of Appendix A for which $c(\mathcal{F}_M; n, r)$ is negative.

M	6–22	33	34	35	38	39	46
r	—	8	15	2	17	10	21
n	—	1, 5, 7	5, 11	1, 5	5, 7, 11	5	5, 7, 11, 13

(The corresponding r-values are marked with a star in the appendix, and a search up to a much larger n-limit yields no further negative coefficients for these values of M.) Here again the asymptotic formula (9.38) helps us to understand this behavior: the negative values are much more likely to occur when κ_r is small. For instance, for $M = 33$ the table

$r\ (= r_{\min})$	1	2	4	5	7	8	10	13	16	19
$\cos(\frac{\pi r}{66}) - \cos(\frac{\pi r^*}{66})$	1.58	1.31	1.71	0.94	1.71	0.49	1.58	1.31	0.94	0.49

shows that the residue classes $r_{\min} \pmod{2M}$ most likely to give some negative Fourier coefficients are 8 and 19, for which the values of κ_r (which are the same; this duplication always occurs for odd M) are minimal, and indeed the class $r = 8$ did lead to three negative coefficients. The class $r = 19$ did not contain any negative values of $c(n, r)$, but instead achieved its small initial average by the initial zeros forced by the optimal growth condition, which says that $c(n, r) = 0$ for $0 \leq n \leq (r^2 - 2)/4m$.

For values of M for more than two prime factors on the other hand, it does not have to be true that $\kappa(\mathcal{F}_M, r_{\min}) > 0$. It is true for the first such index, $M = 210$, and even in a very strong form: not only is κ_r positive for each of the six possible values of $r_{\min}$ (which are nothing other than the values ℓ_0 printed in boldface in Table 4), but of the seven "partners" of each $r_{\min}$, i.e., the seven other values of $r \in [0, 210]$ with $r^2 \equiv r^2_{\min} \pmod{840}$, the three that have the same sign of $C(\Delta, r)$ as for $C(\Delta, r_{\min})$ (i.e., which differ from $r_{\min} \pmod{2p}$ for exactly two of the four prime divisors p of 210) all belong to the first half of the interval $[0, 210]$, while the four that have the opposite sign (i.e., which differ from $r_{\min} \pmod{2p}$ for one or three of these p) belong to the second half, so that all eight terms in the sum in (9.54) are positive. The value of $\kappa(\mathcal{F}_M, r_{\min})$ is again positive for all $r_{\min}$ for the next two values $M = 330$ and $M = 390$, but for $M = 462 = 2\cdot3\cdot7\cdot11$ and $r\ (= r_{\min})\ =\ 31$ we have $\kappa(\mathcal{F}_M, r) < 0$, so that in this case all but finitely many coefficients $C(\mathcal{F}_M;\Delta,r)$ will be negative rather than positive. What does hold for any M and any r (mod $2M$) is that the values of $C(\mathcal{F}_M;\Delta,r)$ eventually all have the same sign, because the coefficient $\kappa(\mathcal{F}_M,r)$ in (9.38) can never vanish, since the primitive $(2M)^{\text{th}}$ roots of unity are linearly independent over $\mathbb{Q}$.

9.6 Higher weights

So far we have studied two families of meromorphic Jacobi forms $\varphi_{1,m}(\tau,z)$ and $\varphi_{2,m}(\tau,z)$ with particularly simple singularities, and the corresponding mock Jacobi forms. We now discuss in much less detail the analogous families for higher weights. We mention that other cases of the decomposition $\varphi=\varphi^F+\varphi^P$ for meromorphic Jacobi forms with higher order poles[27] [17, 98], and the asymptotics of their Fourier coefficients [20], have been studied in the literature.

The starting point is a meromorphic Jacobi form $\varphi(\tau,z)=\varphi_{k,m}(\tau,z)$ of weight k and arbitrary index $m\in\mathbb{N}$ having as its unique singularity (modulo the period lattice $\mathbb{Z}\tau+\mathbb{Z}$) a pole of order k at $z=0$, which we normalized by requiring the leading term of φ to be $(k-1)!/(2\pi iz)^k$. By subtracting a linear combination of functions $\varphi_{k',m}$ with $k'<k$ we can then fix the entire principal part of φ at $z=0$ to be

$$\varphi(\tau,z)\;=\;\frac{(k-1)!}{(2\pi iz)^k}\sum_{0\le 2j<k}\Bigl(\frac{m\,E_2(\tau)}{12}\Bigr)^j\,\frac{(2\pi iz)^{2j}}{j!}\;+\;\mathrm{O}(1)\qquad\text{as } z\to 0.$$

We can make a standard choice, generalizing the previously defined forms (9.2) and (9.1), by taking φ to be the unique polynomial in A^{-1} and B (resp. C times such a polynomial if k is odd) with modular coefficients of the appropriate weight having this principal part, the next three of these being

$$\varphi^{\text{stand}}_{3,m}\;=\;\frac{B^m\,C}{12^m A^2}\,,\qquad\varphi^{\text{stand}}_{4,m}\;=\;\frac{6\,B^{m+2}}{12^{m+2}A^2}\,,\qquad\varphi^{\text{stand}}_{5,m}\;=\;\frac{(B^2-mE_4A^2/2)\,B^{m-1}\,C}{12^m\,A^3}\,.$$

Any other allowed choice of $\varphi_{k,m}$ will differ from this standard one by a weak Jacobi form of weight k and index m, and any two choices will have the same polar part. To write it down, we introduce Euler's functions

$$\mathcal{E}_k(y)\;=\;\mathrm{Av}_{\mathbb{Z}}\Bigl[\frac{(k-1)!}{(2\pi iz)^k}\Bigr]\;=\;\Bigl(-y\,\frac{d}{dy}\Bigr)^{k-1}\mathrm{Av}_{\mathbb{Z}}\Bigl[\frac{1}{2\pi iz}\Bigr]\;=\;\bigl(-y\,\frac{d}{dy}\bigr)^{k-1}\Bigl(\frac{1}{2}\,\frac{y+1}{y-1}\Bigr)\;\in\;\mathbb{Q}\Bigl[\frac{1}{y-1}\Bigr]$$

(which for $k=1$ and $k=2$ coincide with the functions $\mathcal{R}_0$ and $\mathcal{R}_0^{(2)}$ from §8) and their index m averages

$$\mathcal{A}_{k,m}(\tau,z)\;=\;\mathrm{Av}^{(m)}_{\mathbb{Z}\tau+\mathbb{Z}}\Bigl[\frac{(k-1)!}{(2\pi iz)^k}\Bigr]\;=\;\mathrm{Av}^{(m)}\bigl[\mathcal{E}_k(y)\bigr]\;=\;\sum_{s\in\mathbb{Z}}q^{ms^2}y^{2ms}\mathcal{E}_k(q^sy)$$

---

[27] These were considered in [17] while answering a question due to Kac concerning the modularity of certain characters of affine Lie superalgebras.

generalizing (9.3). These functions all have wall-crossing, with Fourier expansions given by

$$\mathcal{E}_k(y) \;=\; \begin{cases} \frac{1}{2}\,\delta_{k,1} \,+\, \sum_{t>0} t^{k-1}y^{-t} & \text{if } |y|>1 \\ -\frac{1}{2}\,\delta_{k,1} \,-\, \sum_{t<0} t^{k-1}y^{-t} & \text{if } |y|<1 \end{cases} \;=\; \sum_{t\in\mathbb{Z}} \frac{\operatorname{sgn}(t)-\operatorname{sgn}(z_2)}{2}\, t^{k-1}y^{-t}$$

and

$$\mathcal{A}_{k,m}(\tau,z) \;=\; \sum_{s,\,t\in\mathbb{Z}} \frac{\operatorname{sgn}(t) \,-\, \operatorname{sgn}(s+z_2/\tau_2)}{2}\, t^{k-1}q^{ms^2-st}y^{2ms-t}\,. \tag{9.55}$$

(Note that in the latter expression the coefficient of y^r is a Laurent polynomial in q for every r and that the discriminant $4nm-r^2$ is the negative of a perfect square for each monomial q^ny^r occurring.) Then the polar part of $\varphi_{k,m}$ is given by

$$\varphi^P_{k,m}(\tau,z) \;=\; \sum_{0\le 2j<k} \frac{(k-1)!}{(k-1-2j)!\,j!}\Bigl(\frac{m\,E_2(\tau)}{12}\Bigr)^j\,\mathcal{A}_{k-2j,m}(\tau,z)\,.$$

(To see this, observe that by (9.55) each of the functions $\mathcal{A}_{k,m}$, and hence any linear combination of them with coefficients depending only on τ, is in the kernel of the operator $\varphi\mapsto h_\ell$ defined by (8.2) for all $\ell\in\mathbb{Z}$.) Its finite part $\Phi_{k,m} = \varphi^F_{k,m} = \varphi_{k,m}-\varphi^P_{k,m}$ is a weakly holomoprhic elliptic form that depends on the particular choice of $\varphi_{k,m}$. From the results of §10 (specifically, Theorem 10.1) it follows that for $k\ge 3$ the function $\varphi_{k,m}$ can always be chosen so that $\Phi_{k,m}$ is strongly holomorphic, this choice then obviously being unique up to the addition of a strongly holomorphic Jacobi form.

As an example, consider $k=3$ and $m=2$. (This is the smallest possibility since index 1 cannot occur in odd weight.) The finite part of our standard choice is

$$\Phi^{\text{stand}}_{3,2} \;=\; \frac{B^2C}{144A^2} \,-\, \mathcal{A}_{3,2} \,-\, \frac{E_2}{3}\,\mathcal{A}_{3,1} \;=\; h^{\text{stand}}_{3,2}\,\bigl(\vartheta_{2,1} \,-\, \vartheta_{2,3}\bigr)$$

with

$$h^{\text{stand}}_{3,2} \;=\; \frac{1}{144}\,q^{-1/8}\,\bigl(1+51q+2121q^2+14422q^3+60339q^4+201996q^5+588989q^6+\cdots\bigr)\,.$$

This is weakly holomorphic and has large Fourier coefficients, but by subtracting a multiple of the unique weak Jacobi form E_4C of this weight and index we get a new choice

$$\Phi^{\text{opt}}_{3,2} \;=\; \Phi^{\text{stand}}_{3,2} \,-\, \frac{E_4}{144}\,C \;=\; h^{\text{opt}}_{3,2}\,\bigl(\vartheta_{2,1} \,-\, \vartheta_{2,3}\bigr)$$

that is strongly holomorphic and has much smaller coefficients

$$h^{\text{opt}}_{3,2} \;=\; h^{\text{stand}}_{3,2} \,-\, \frac{E_4}{144\,\eta^3} \;=\; -\frac{4}{3}\,q^{-1/8}\,\bigl(q+4q^2+5q^3+11q^4+6q^5+21q^6+\cdots\bigr)\,,$$

in accordance with the theorem just cited.

However, this choice, although optimal from the point of view of the order of growth of its Fourier coefficients, is not the most enlightening one. By subtracting a different multiple of E_4C from $\Phi_{3,2}^{\text{stand}}$, we get a third function

$$\Phi_{3,2} \;=\; \Phi_{3,2}^{\text{stand}} + \frac{7E_4}{144}\,C \;=\; \Phi_{3,2}^{\text{opt}} + \frac{E_4}{18}\,C \;=\; h_{3,2}\bigl(\vartheta_{2,1} \,-\, \vartheta_{2,3}\bigr)$$

with

$$h_{3,2} \;=\; h_{3,2}^{\text{stand}} + \frac{7E_4}{144\,\eta^3} \;=\; h_{3,2}^{\text{opt}} + \frac{E_4}{18\,\eta^3} \;=\; \frac{1}{18}\,q^{-1/8}\,(1 + 219q + 2793q^2 + 15262q^3 + \cdots)\,.$$

This function can be identified: it is given by

$$h_{3,2} \;=\; \frac{1}{3}\,\partial_{1/2}^{\text{RS}}(h_2)\,,$$

where $h_2(\tau) = q^{-1/8}(-1 + 45q + 231q^2 + \cdots)$ is the mock theta function defined in (7.16) (the one occurring in the Mathieu moonshine story) and $\partial_{1/2}^{\text{RS}}$ is the Ramanujan-Serre derivative as defined in (3.7). We can also write this at the level of mock Jacobi forms as

$$\Phi_{3,2} \;=\; \frac{1}{24}\,\mathcal{L}_{1,2}(\mathcal{F}_2)\,,$$

where $\mathcal{F}_2$ is the mock Jacobi form defined in (7.32) and $\mathcal{L}_{1,2}$ the modified heat operator defined in (4.12).

An analogous statement holds in general, e.g., for weight 3 we have

$$\Phi_{3,m} \;=\; -\,\mathcal{L}_{1,m}\bigl(\Phi_{1,m}\bigr) \qquad \pmod{\widetilde{J}_{3,m}}$$

for any index m and any choice of the functions $\Phi_{3,m}$ and $\Phi_{1,m}$ in their defining classes, and similarly for higher weight, the specific formulas for $k \in \{3,4\}$ and the standard choices being

$$\begin{aligned}
\Phi_{3,m}^{\text{stand}} + \mathcal{L}_{1,m}\bigl(\Phi_{1,m}^{\text{stand}}\bigr) \;&=\; -\,\frac{m-1}{12^m}\Bigl(4(m-2)E_6A \,+\, (2m+3)E_4B\Bigr)\,B^{m-3}C\,,\\
\Phi_{4,m}^{\text{stand}} + \mathcal{L}_{2,m}\bigl(\Phi_{2,m}^{\text{stand}}\bigr) \;&=\; -\,\frac{m+1}{72\cdot 12^m}\Bigl(4mE_6A \,+\, (2m-3)E_4B\Bigr)\,B^{m-1}\,.
\end{aligned}$$

Thus in one sense, the higher-weight families are better than the weight 1 and weight 2 families (because they can be made strongly holomorphic by subtracting an appropriate weak Jacobi form), but in another, they are much less interesting (because they are always combinations of derivatives of lower-weight examples and ordinary Jacobi forms, so that they do not produce any essentially new functions).

10. Structure theorems for Jacobi and mock Jacobi forms

In the last section we introduced two special families of meromorphic Jacobi forms and the associated mock Jacobi forms and described various observations concerning them. In this section we try to give theoretical explanations for some of these. We begin by studying in §10.1 how the holomorphic Jacobi forms of weight k and index m lie in the space of weak Jacobi forms of the same weight and index, or equivalently, what the relations are among the polar Fourier coefficients (i.e., the coefficients of $q^n y^r$ with $4nm < r^2$) of weak Jacobi forms. The results described here are relatively easy consequence of results in the literature but are, as far as we know, new. The remaining two subsections describe the application of these results to mock Jacobi forms in general and to our two special families in particular, providing fairly complete explanations of the experimentally found properties of the weight two family and partial explanations in the weight one case.

10.1 Relations among the polar coefficients of weak Jacobi forms

The definitions and main properties of holomorphic and weak Jacobi forms were reviewed in §4, following the more detailed treatment in [57]. Here we investigate the relationship of these two spaces in more detail.

By definition, a weak Jacobi form has coefficients $c(n,r)$ which are non-zero only for $n \geq 0$, whereas a holomorphic one has coefficients that vanish unless $4nm \geq r^2$. By virtue of the periodicity property (4.4), this stronger condition need only be checked for r running over a set of representatives for $\mathbb{Z}/2m\mathbb{Z}$, say $-m < r \leq r$, and since condition (4.1) with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ also implies that $c(n,-r) = (-1)^k c(n,r)$ and consequently $C(\Delta,-r) = (-1)^k C(\Delta,r)$, it in fact suffices to consider the coefficients with $0 \leq r \leq m$ if k is even and those with $0 < r < m$ if k is odd. In other words, for all positive integers k and m we have an exact sequence

$$0 \longrightarrow J_{k,m} \longrightarrow \widetilde{J}_{k,m} \xrightarrow{P} \mathbb{C}^{\mathcal{N}_{\pm,m}} \qquad (\,(-1)^k = \pm 1\,)\,, \tag{10.1}$$

where

$$\begin{aligned} \mathcal{N}_{+,m} &= \left\{(r,n) \in \mathbb{Z}^2 \;\middle|\; 0 \leq r \leq m,\ 0 \leq n < r^2/4m\right\}\,, \\ \mathcal{N}_{-,m} &= \left\{(r,n) \in \mathbb{Z}^2 \;\middle|\; 0 < r < m,\ 0 \leq n < r^2/4m\right\}\,, \end{aligned}$$

and where the map P ("polar coefficients") associates to $\varphi \in \widetilde{J}_{k,m}$ the collection of its Fourier coefficients $c_\varphi(n,r)$ for $(n,r) \in \mathcal{N}_{\pm,m}$. In particular we have

$$\dim J_{k,m} \;\geq\; j(k,m) \;:=\; \dim \widetilde{J}_{k,m} \;-\; \left|\mathcal{N}_{\pm,m}\right| \qquad (\,(-1)^k \;=\; \pm 1\,)\,, \tag{10.2}$$

with equality if and only if the map P in (10.1) is surjective. (Cf. [57], p. 105.)

Using the description of $\widetilde{J}_{*,*}$ given in §4 (eqs. (4.25) and (4.35)) and the definition of $\mathcal{N}_{\pm,m}$, we can write this lower bound on the dimension explicitly as

$$j(k,m) \;=\; \begin{cases} \sum_{0\le j\le m}\big(\dim M_{k+2j} - \lceil j^2/4m\rceil\big) & \text{if } k \text{ is even,} \\ \sum_{0< j< m}\big(\dim M_{k+2j-1} - \lceil j^2/4m\rceil\big) & \text{if } k \text{ is odd,} \end{cases} \tag{10.3}$$

where $\lceil x\rceil$ denotes the "ceiling" of a real number x ($=$ smallest integer $\ge x$). This formula is computable and elementary since $\dim M_k = k/12 + \mathrm{O}(1)$ with the $\mathrm{O}(1)$ term depending only on $k \pmod{12}$, but is nevertheless not a good formula because the two terms $\sum \dim M_{k+2j}$ and $\sum \lceil j^2/4m\rceil$ are each equal to $m^2/12 + \mathrm{O}(m)$, and $j(k,m)$ is only of the order of m. In [57], pp. 122–125 it was shown that, somewhat surprisingly, the numbers $j(k,m)$ can be given by a completely different-looking formula involving class numbers of imaginary quadratic fields. This formula has the form

$$j(k,m) \;=\; \frac{2k-3}{24}\,(m+1) \,+\, \frac{b}{4} \,-\, \frac{1}{4}\sum_{d|4m} h(-d) \,+\, \mathrm{O}(1)\,, \tag{10.4}$$

where $m = ab^2$ with a square-free and where the $\mathrm{O}(1)$ term depends only on m and k modulo 12. In particular, $j(k,m) = \frac{2k-3}{24}\,m \,+\, \mathrm{O}(m^{1/2+\varepsilon})$ as $m\to\infty$ with k fixed. The formula (10.4) has an interesting interpretation as the dimension of a certain space of classical modular forms of level m and weight $2k-2$, discussed in detail on pp. 125–131 of [57].

In [57], it was shown that the inequality (10.2) is an equality for $k \ge m$ and it was conjectured that this holds for all $k \ge 3$. This was later proved in [115] as a consequence of a difficult trace formula calculation. In fact, a slightly stronger statement was proved there: It is clear that the exact sequence (10.1) remains valid if one replaces $J_{k,m}$ by $J^0_{k,m}$ and the map $P : \widetilde{J}_{k,m} \to \mathbb{C}^{\mathcal{N}_{\pm,m}}$ by the analogously defined map $P^0 : \widetilde{J}_{k,m} \to \mathbb{C}^{\mathcal{N}^0_{\pm,m}}$, where $\mathcal{N}^0_{\pm,m}$ is defined exactly like $\mathcal{N}_{\pm,m}$ but with the strict inequality $n < r^2/4m$ replaced by $n \le r^2/4m$. Then the dimension formula proved in [115] says that $\dim J^0_{k,m} = \dim \widetilde{J}_{k,m} - |\mathcal{N}^0_{\pm,m}|$ and hence that the map P^0 is surjective, which is a stronger statement than the surjectivity of P since the set $\mathcal{N}^0_{\pm,m}$ is larger than $\mathcal{N}_{\pm,m}$. In summary, the results of [57] and [115] together give the following complete description of the "polar coefficients" maps for both holomorphic and cuspidal Jacobi forms when $k \ge 3$:

Theorem 10.1. *For $k > 2$ and all $m \ge 1$ the map $P^0 : \widetilde{J}_{k,m} \to \mathbb{C}^{\mathcal{N}^0_{\pm,m}}$ is surjective and we have a commutative diagram of short exact sequences*

$$\begin{tikzcd}
0 \arrow[r] & J^0_{k,m} \arrow[r] \arrow[d, phantom, "\cap"] & \widetilde{J}_{k,m} \arrow[r, "P^0"] \arrow[d, equal] & \mathbb{C}^{\mathcal{N}^0_{\pm,m}} \arrow[r] \arrow[d, twoheadrightarrow] & 0 \\
0 \arrow[r] & J_{k,m} \arrow[r] & \widetilde{J}_{k,m} \arrow[r, "P"] & \mathbb{C}^{\mathcal{N}_{\pm,m}} \arrow[r] & 0
\end{tikzcd} \tag{10.5}$$

We now turn to the case $k=2$, the one of main interest for us. Here formula (10.4) takes the explicit form

$$j(2,m) \;=\; \frac{1}{2}\Bigl(\Bigl\lfloor\frac{m+1}{3}\Bigr\rfloor \,+\, \Bigl\lfloor\frac{m+1}{3}\Bigr\rfloor\Bigr) \,-\, \frac{m-b-1}{4} \,-\, \frac{1}{4}\sum_{d|4m} h(-d)\,, \tag{10.6}$$

but for this value of k it was observed in [57] (p. 131) that equality could not always hold in (10.2) (indeed, the right-hand side is often negative) and that, for reasons explained in detail there, one should replace the number $j(k,m)$ by the modified number

$$j^*(2,m) \;:=\; j(2,m) \,+\, \delta(m)\,, \qquad \text{where} \qquad \delta(m) \;:=\; \sum_{\substack{d\le\sqrt{m}\\ d|m,\; d^2\nmid m}} 1\;. \tag{10.7}$$

We note that $\delta(m)$ is a linear combination of three multiplicative functions of m, namely

$$\delta(m) \;=\; \frac{1}{2}\sigma_0(m) \,+\, \frac{1}{2}\,\delta_{a,1} \,-\, \sigma_0(b) \tag{10.8}$$

(with $m=ab^2$, a square-free as before), or equivalently in terms of a generating Dirichlet series

$$\sum_{m=1}^{\infty}\frac{\delta(m)}{m^s} \;=\; \frac{1}{2}\,\zeta(s)^2 \,+\, \frac{1}{2}\,\zeta(2s) \,-\, \zeta(s)\zeta(2s)\;. \tag{10.9}$$

We also note that $\delta(m)$ depends only on the "shape" of m (= the collection of exponents ν_i if m is written as $\prod p_i^{\nu_i}$), the first three cases being

$$\delta(p^\nu) \;=\; 0\,, \tag{10.10}$$

$$\delta(p_1^{\nu_1}p_2^{\nu_2}) \;=\; n_1n_2\,, \tag{10.11}$$

$$\delta(p_1^{\nu_1}p_2^{\nu_2}p_3^{\nu_3}) \;=\; 3n_1n_2n_3+\varepsilon_1n_2n_3+\varepsilon_2n_1n_3+\varepsilon_3n_1n_2\,, \tag{10.12}$$

in which we have written each ν_i as $2n_i+\varepsilon_i-1$ with $n_i\ge1$ and $\varepsilon_i\in\{0,1\}$.

The conjecture made in [57] that $\dim J_{k,m}=j^*(2,m)$ for all $m\ge1$ was proved in [115], but here we want to give a sharper statement describing the cokernel of the map P explicitly. To do this, we follow [57], pp. 132, where an elementary argument (containing a small error which will be corrected below) was given to show that $\dim J_{2,m}\ge j^*(2,m)$ for square-free m. Actually, it is more convenient to work with cusp forms, since the corresponding statement here is

$$\dim J^0_{2,m} \;=\; \dim \widetilde{J}_{2,m} \,-\, |\mathcal{N}^0_{+,m}| \,+\, \delta^0(m) \tag{10.13}$$

with

$$\delta^0(m) \;=\; \sum_{d|m,\; d\le\sqrt{m}} 1 \;=\; \sum_{\substack{m=m_1\cdot m_2\\ \text{up to order}}} 1\,, \tag{10.14}$$

and $\delta^0(m)$ is a slightly simpler function than $\delta(m)$. (The third term in (10.8) is missing.) We want to explain formula (10.13) by finding an explicit map

$$\beta\,:\,\mathbb{C}^{\mathcal{N}^0_{+,m}}\;=\;\bigoplus_{\substack{0\le r\le m\\ 0\le n\le r^2/4m}}\mathbb{C}\;\longrightarrow\;\Delta^0(m):=\bigoplus_{\substack{m=m_1\cdot m_2\\ \text{up to order}}}\mathbb{C} \tag{10.15}$$

such that the sequence

$$0\longrightarrow J^0_{2,m}\longrightarrow \widetilde{J}_{2,m}\xrightarrow{P^0}\mathbb{C}^{\mathcal{N}^0_{+,m}}\xrightarrow{\beta}\Delta^0(m)\longrightarrow 0\;. \tag{10.16}$$

is exact.

To give a natural definition of the map β, we work, not with $\widetilde{J}_{2,m}$ or even with $J^!_{2,m}$, but with the larger vector spaces $\widetilde{\mathcal{E}}_{+,m}$ and $\mathcal{E}^!_{+,m}$ of even weak or weakly holomorphic elliptic forms of index m, as defined at the end of §4.4. We define a "residue map" $\mathcal{R}:\mathcal{E}^!_{+,m}\to\mathbb{C}$ by

$$\varphi\;=\;\sum_{n\ge0,\,r\in\mathbb{Z}}c(n,r)\,q^n\,y^r\quad\mapsto\quad\mathcal{R}[\varphi]\;=\;\mathrm{Res}_{\tau=\infty}\big(\varphi(\tau,0)\,d\tau\big)\;=\;\sum_{r\in\mathbb{Z}}c(0,r)\;. \tag{10.17}$$

Then $\mathcal{R}[\varphi]=0$ if $\varphi\in J^!_{2,m}$ because in that case $\varphi(\,\cdot\,,0)$ belongs to $M^!_2$ and the q^0 term of any weakly holomorphic modular form f of weight two is zero by the residue theorem applied to the $SL(2,\mathbb{Z})$-invariant 1-form $f(\tau)d\tau$. (The map $\mathcal{R}$ was already used in [57], but it was erroneously stated there that for $\varphi\in\widetilde{J}_{2,m}$ the function $\varphi(\tau,0)$ itself, rather than just its residue at infinity, vanishes.) Now for each decomposition $m=m_1\cdot m_2$ (up to order, i.e., without distinguishing the splittings $m_1\cdot m_2$ and $m_2\cdot m_1$) we define $R_{m_1,m_2}:\mathcal{E}^!_{+,m}\to\mathbb{C}$ by setting $R_{m_1,m_2}(\varphi)=\mathcal{R}\big[\varphi\big|W_{m_1}\big]$ if $(m_1,m_2)=1$, where W_{m_1} is the involution defined in §4.4, while if $(m_1,m_2)=t>1$ then we define $R_{m_1,m_2}(\varphi)$ as $R_{m_1/t,m_2/t}(\varphi|u_t)$, where u_t is the map from $\mathcal{E}^!_{+,m}$ to $\mathcal{E}^!_{+,m/t^2}$ defined in §4.4. (It was pointed out there that both W_{m_1} for $(m_1,m/m_1)=1$ and u_t for $t^2|m$ make sense on all of $\mathcal{E}^!_{+,m}$, and not just on the subspace of holomorphic Jacobi forms, but that this is *not* true for weak elliptic or weak Jacobi forms, which is why we are now working with weakly holomorphic forms.) We put all of these maps together to a single map $R:\,\mathcal{E}^!_{+,m}\to\Delta^0(m)$ and then define the desired map β by the commutative diagram

$$\begin{tikzcd}
\widetilde{\mathcal{E}}_{+,m} \arrow[rr, "P^0"] \arrow[d, phantom, "\cap"] & & \mathbb{C}^{\mathcal{N}^0_{+,m}} \arrow[d, phantom, "\cap"] \\
\mathcal{E}^!_{+,m} \arrow[rr] \arrow[dr, "R"'] & & \mathbb{C}^{\mathcal{N}^{!,0}_{+,m}} \arrow[dl, "\beta"] \\
& \Delta^0(m) &
\end{tikzcd} \tag{10.18}$$

with $\mathbb{C}^{\mathcal{N}^{!,0}_{+,m}}$ defined as in (10.23) below. (To see that such a map β exists, observe that the coefficients $c(0,r)$ occurring in (10.17) all have discriminant $\Delta=-r^2\le0$, that the operators

W_{m_1} and u_t can be defined purely in terms of the coefficients $C_\varphi(\Delta,r)$ and preserve the condition $\Delta\le 0$, and finally that the coefficients $C_\varphi(\Delta,r)$ for $\varphi\in\widetilde{\mathcal{E}}_{+,m}$ can all be expressed in terms of the coefficients $c_\varphi(n,r)$ with $(n,r)\in\mathcal{N}_{+,m}$ by virtue of the periodicity and evenness conditions defining $\widetilde{\mathcal{E}}_{+,m}$.) The discussion above shows that the composition $\beta\circ P^0$ in (10.16) vanishes, since $\mathcal{R}$ vanishes on weakly holomorphic and hence in particular on weak Jacobi forms, and combining this statement with the dimension formula (10.13) proved in [115] we obtain the main result describing strong and weak Jacobi forms of weight 2 and arbitrary index:

Theorem 10.2. *The sequence* (10.16) *with* β *defined by* (10.18) *is exact for all integers* $m\ge 1$.

The two theorems above give a complete description of the space of Jacobi forms of all weights $k\ge 2$. We end this subsection by discussing the case of weight one. Here the description of Jacobi forms is very simple to state (though its proof is quite deep), since we have the following theorem of Skoruppa ([111], quoted as Theorem 5.7 of [57]):

Theorem 10.3. *We have* $J_{1,m}=\{0\}$ *for all* $m\ge 1$.

Combining this with the discussion above, we find that in this case we have an exact sequence

$$0\longrightarrow \widetilde{J}_{1,m}\longrightarrow \bigoplus_{\substack{0<r<m\\ 0\le n<r^2/4m}}\mathbb{C}\longrightarrow \mathbb{C}^{|j(1,m)|}\longrightarrow 0\,. \tag{10.19}$$

with

$$|j(1,m)| \;=\; -j(1,m)\;=\;\frac{m}{24}\;+\;\mathrm{O}\big(m^{1/2+\varepsilon}\big)\,.$$

We have not been able to find a natural description of the last space, or explicit description of the last map, in (10.19), analogous to the results given for $k=2$.

10.2 Choosing optimal versions of weak elliptic forms

In this subsection, we will apply the theorem just explained to show how to make "optimal versions" of elliptic forms by counting actual Jacobi forms to get poles of small order.

Let $\widetilde{\mathcal{E}}_{+,m}$ be as in the previous subsection and $\widetilde{\mathcal{E}}_{-,m}$ be the similarly defined space with $c(n,-r)=-c(n,r)$, and let $\mathcal{E}_{\pm,m}$ and $\mathcal{E}^0_{\pm,m}$ be the corresponding spaces with the condition (4.7) replaced by (4.5) and (4.6), respectively. The discussion of §10.1 can be summarized as saying that we have a commutative diagram of short exact sequences,

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{E}_{\pm,m} \arrow[r] & \widetilde{\mathcal{E}}_{\pm,m} \arrow[r, "P"] & \mathbb{C}^{\mathcal{N}_{\pm,m}} \arrow[r] & 0 \\
0 \arrow[r] & J_{k,m} \arrow[r] \arrow[u, phantom, "\cup"] & \widetilde{J}_{k,m} \arrow[r] \arrow[u, phantom, "\cup"] & \mathbb{C}^{\mathcal{N}_{\pm,m}} \arrow[r] \arrow[u, phantom, "||"] & 0
\end{tikzcd} \tag{10.20}$$

for $k > 2$, $(-1)^k = \pm 1$, and similarly with $J_{k,m}$ and $\mathcal{E}_{\pm,m}$ replaced by $J^0_{k,m}$ and $\mathcal{E}^0_{\pm,m}$, $\mathcal{N}_{\pm,m}$ by $\mathcal{N}^0_{\pm,m}$ and P by P^0. For $k = 2$, we have instead:

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{E}^0_{+,m} \arrow[r] \arrow[d, phantom, "\cup"] & \widetilde{\mathcal{E}}_{+,m} \arrow[r, "P^0"] \arrow[d, phantom, "\cup"] & \mathbb{C}^{\mathcal{N}^0_{+,m}} \arrow[r] \arrow[d, equal] & 0 & \\
0 \arrow[r] & J^0_{2,m} \arrow[r] & \widetilde{J}_{2,m} \arrow[r] & \mathbb{C}^{\mathcal{N}^0_{+,m}} \arrow[r, "\beta"] & \Delta^0(m) \arrow[r] & 0
\end{tikzcd} \tag{10.21}$$

in the cuspidal case, where $\Delta^0(m)$ was defined in (10.15) and β is defined by $\beta \circ P^0 = R$. As a consequence of these diagrams and of the vanishing of $\delta(p^\nu)$ (*cf.* equation (10.10)), which implies that the diagram (10.20) also holds for $k = 2$ when m is a prime power, we obtain:

Theorem 10.4. *Let $k \geq 2$, and $\varphi \in \widetilde{\mathcal{E}}_{\pm,m}$, where $(-1)^k = \pm 1$. Then φ is the sum of an element of $\widetilde{J}_{k,m}$ and an element of $\mathcal{E}^0_{\pm,m}$ if and only if either*

(i) $\quad k > 2\,,$ *or*

(ii) $\quad k = 2$ *and* $R(\varphi)\ (= \beta(P^0(\varphi)))\ = 0\,;$

If $k = 2$ and m is a prime power, then φ is the sum of an element of $\widetilde{J}_{k,m}$ and an element of $\mathcal{E}_{\pm,m}$.

Theorem 10.4 for $k = 2$ makes a negative statement: if $\varphi \in \widetilde{\mathcal{E}}_{+,m}$ has $\mathcal{R}[\varphi] \neq 0$, and m is not a prime power, then φ *cannot* in general be decomposed as the sum of a weak Jacobi form of weight 2 and a holomorphic elliptic form. However, it turns out that we can come very close to this, writing φ as a sum of images under Hecke-like operators of forms that "have optimal growth" in the sense of §9.3. We now explain this in two steps. We first use the decomposition (4.44), and its variants for weakly holomorphic Jacobi or elliptic forms, to reduce to the case of primitive forms. We then show that any primitive even elliptic form can be corrected by the addition of a weakly holomorphic weight 2 Jacobi form to give a new elliptic form whose non-zero Fourier coefficients $c(n, r)$ all have discriminant $4mn - r^2 \geq -1$.

Since the decomposition (4.44), as was already observed in §4.3 fails for weak Jacobi forms but holds for weakly holomorphic ones, we must first extend the above results to weakly holomorphic forms.

Proposition 10.1. *All statements in Theorem 10.4 remain true* mutatis mutandis *if all $\widetilde{\mathcal{E}}$, $\widetilde{J}$ are replaced by $\mathcal{E}^!$, $J^!$.*

The proposition is equivalent to saying that the analog of (10.21) still holds with weakly holomorphic rather than weak forms, i.e. we have a diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{E}^0_{+,m} \arrow[r] \arrow[d, phantom, "\cup"] & \mathcal{E}^!_{+,m} \arrow[r, "P^0"] \arrow[d, phantom, "\cup"] & \mathbb{C}^{\mathcal{N}^{0,!}_{+,m}} \arrow[r] \arrow[d, equal] & 0 & \\
0 \arrow[r] & J^0_{2,m} \arrow[r] & J^!_{2,m} \arrow[r] & \mathbb{C}^{\mathcal{N}^{0,!}_{+,m}} \arrow[r, "\beta"] & \Delta^0(m) \arrow[r] & 0
\end{tikzcd} \tag{10.22}$$

where

$$\mathcal{N}^{0,!}_{+,m} \;=\; \bigl\{(r,n)\in\mathbb{Z}^2 \;\mid\; 0\le r\le m,\; -\infty<n\le r^2/4m\bigr\}\,. \tag{10.23}$$

We again prove this by comparing dimensions, even though the middle spaces in each row in the diagram (10.22) are infinite-dimensional. Since any elliptic form φ has a minimal value of n with $c_\varphi(n,r)\neq 0$ for some r, it suffices to prove the exactness of the rows in (10.22) with $\mathcal{E}^!_{+,m}$, $J^!_{2,m}$, and $\mathbb{C}^{\mathcal{N}^{0,!}_{+,m}}$ replaced by the spaces $\mathcal{E}^{\ge N}_{+,m}$, $J^{\ge N}_{2,m}$, and $\mathbb{C}^{\mathcal{N}^{0,\ge N}_{+,m}}$, of forms with non-zero coefficients only for $n\ge -N$ for some $N\ge 0$. The only non-trivial part is the exactness at $\mathbb{C}^{\mathcal{N}^{0,\ge -N}_{+,m}}$, and since the composition of the maps P^0 and β vanishes, we only need to check that the alternating sum of dimensions is zero. For this, we observe that the dimensions of $J^{\ge -N}_{2,m}$ and $\mathbb{C}^{\mathcal{N}^{0,\ge -N}_{+,m}}$ are both larger by $(m+1)N$ than their values at $N=0$: for $\mathbb{C}^{\mathcal{N}^{0,\ge -N}_{+,m}}$, this is obvious since $\mathcal{N}^{0,\ge -N}_{+,m}$ differs from $\mathcal{N}^0_{+,m}$ by the addition of N rows of length $m+1$, and for $J^{\ge -N}_{2,m}$ it follows from the computation

$$\begin{aligned}\dim J^{\ge -N}_{2,m} \;&=\; \dim \widetilde{J}_{12N+2,m} \qquad \bigl(\text{because } J^{\ge -N}_{2,m} = \Delta(\tau)^{-N}\widetilde{J}_{12N+2,m}\bigr)\\ &=\; \sum_{j=0}^{m}\dim M_{k+12N+2j} \;=\; \dim \widetilde{J}_{2,m} \;+\; (m+1)N\,. \end{aligned} \tag{10.24}$$

This completes the proof of the proposition. □

Now we can apply the decomposition (4.44) to the whole diagram (10.22) to find that everything can be reduced to primitive forms (and their images under the U_t operators). The key point is that although U_t and u_t change the discriminant $\Delta = 4mn - r^2$ by multiplying or dividing it by t^2, respectively, they do not change its sign, and therefore all of the spaces in our diagram can be decomposed as direct sums of primitive parts of lower index and their images under U_t. This reduces everything to case of primitive forms.

For primitive forms our diagram becomes

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{E}^{0,\mathrm{prim}}_{+,m} \arrow[r] & \mathcal{E}^{!,\mathrm{prim}}_{+,m} \arrow[r, "P^0"] & \mathbb{C}^{\mathcal{N}^{0,!,\mathrm{prim}}_{+,m}} \arrow[r] & 0 & \\ & \cup & \cup & \| & & \\ 0 \arrow[r] & J^{0,\mathrm{prim}}_{2,m} \arrow[r] & J^{!,\mathrm{prim}}_{2,m} \arrow[r] & \mathbb{C}^{\mathcal{N}^{0,!,\mathrm{prim}}_{+,m}} \arrow[r, "\beta"] & \Delta^{0,\mathrm{prim}}(m) \arrow[r] & 0 \end{tikzcd} \tag{10.25}$$

where

$$\Delta^{0,\mathrm{prim}}(m) = \bigoplus_{\substack{m=m_1\cdot m_2\\ (m_1,m_2)=1\\ \text{up to order}}} \mathbb{C}\,. \tag{10.26}$$

The map β in (10.25) is again defined by the requirement that $\beta\circ P^0 = R$ on $\mathcal{E}^{!,\mathrm{prim}}_{+,m}$, and lands in $\Delta^{0,\mathrm{prim}}(m)$ because $\varphi|u_t = 0$ for φ primitive and $t>1$. The space $\Delta^{0,\mathrm{prim}}(m)$ has dimension $\delta^{0,\mathrm{prim}}(m) = 2^{s-1}$ if $m>1$, where s is the number of primes dividing m. (The failure of this formula for $m=1$ corresponds to the second term in (10.8).)

We also have the non-cusp-form analogs of the diagrams (10.22) and (10.25), where all superscripts "0" have been dropped and where $\mathcal{N}^!_{+,m}$ is defined as in (10.23) but with the condition $4mn - r^2 \leq 0$ replaced by the strict inequality $4mn - r^2 < 0$, while $\Delta^{\rm prim}(m)$ is a space of dimension $\delta^{\rm prim}(m) = \delta^{0,{\rm prim}}(m) - 1 = 2^{s-1} - 1$ (compare equation (10.8)). This space is the quotient of the space in (10.26) by the one-dimensional subspace of vectors with all components equal. This is essentially the idea of the proof of Theorem 9.4 given below.

Having completed these preparations, we now proceed to the proofs of our main results. We consider the case of square-free index first and then give a brief description of the modifications needed to extend them to primitive forms of arbitrary index.

For $m > 1$ square-free, we will show that any $\varphi \in \widetilde{\mathcal{E}}_{+,m}$ can be corrected by a weak Jacobi form of weight 2 to get a another elliptic form whose non-zero Fourier coefficients $c(n, r)$ all have discriminant $4mn - r^2 \geq -1$, i.e., that we have a decomposition

$$\widetilde{\mathcal{E}}_{+,m} \;=\; \widetilde{J}_{2,m} + \mathcal{E}^{\rm OG}_{+,m} \qquad (m \text{ square-free}), \tag{10.27}$$

with $\mathcal{E}^{\rm OG}_{+,m}$ defined as in §9.3. Actually, this is still not quite the statement we want, since it is slightly wasteful: the codimension of $\mathcal{E}^0_{+,m}$ in $\mathcal{E}^{\rm OG}_{+,m}$ is $2^{s-1}+1$, where s as before is the number of prime factors of m (there are 2^{s-1} pairs $(r, n) \in \mathcal{N}^0_{+,m}$ with $\Delta := 4nm - r^2$ equal to -1, and one pair (0,0) with $\Delta = 0$), and there are only $\dim \Delta^0(m) = 2^{s-1}$ constraints $R_{m_1,m_2}(\varphi) = 0$ coming from diagram (10.21). Correspondingly, the intersection of the two spaces on the right-hand side of (10.27) contains $J_{2,m}$ ($= J^0_{2,m}$ in this case) as a subspace of codimension 1. To get a sharp statement, we should therefore replace $\mathcal{E}^{\rm OG}_{+,m}$ by a suitable codimension one subspace. There are three natural ways to do this, each of which leads to a sharpening of (10.27), corresponding exactly (in the square-free case) to the various choices of $\Phi^0_{2,m}$ introduced in the discussion preceding Theorem 9.5. What's more, all statements remain true verbatim for arbitrary $m > 1$ if we restrict to primitive forms. We state all three variants in a single theorem.

Theorem 10.5. *Suppose that $m > 1$ is square-free and let $\mathcal{E}^{{\rm OG},\star}_m$ ($\star$ = I, II, III) denote the codimension one subspace of $\mathcal{E}^{\rm OG}_{+,m}$ defined by one of the three conditions*

(I) $\sum_{0<r<m,\; r^2\equiv 1 \pmod{4m}} C_\varphi(-1; r) = 0$,

(II) $c_\varphi(0, 1) = 0$,

(III) $c_\varphi(0, 0) = 0$,

respectively. Then we have an exact sequence

$$0 \longrightarrow J_{2,m} \xrightarrow{\rm diag} \widetilde{J}_{2,m} \oplus \mathcal{E}^{{\rm OG},\star}_m \xrightarrow{-} \widetilde{\mathcal{E}}_{+,m} \longrightarrow 0\;, \tag{10.28}$$

i.e., any even weak elliptic form of index m can be decomposed as the sum of a weak Jacobi form of weight 2 and index m and an element of $\mathcal{E}^{{\rm OG},\star}_m$, and this decomposition is unique up to a holomorphic Jacobi form (which can be added to one term and subtracted from the other).

For $m > 1$ the same statement remains true if we replace $\widetilde{\mathcal{E}}_{+,m}$ by its subspace $\widetilde{\mathcal{E}}^{\mathrm{res-prim}}_{+,m}$ of forms φ satisfying $R(\varphi|u_t) = 0$ for all $t > 1$ and $\mathcal{E}^{\mathrm{OG},\star}_{m}$ by its intersection with $\widetilde{\mathcal{E}}^{\mathrm{res-prim}}_{+,m}$.

Proof. By the exactness of the top line of diagram (10.21), we can find functions $\psi_{m;(r,n)} \in \widetilde{\mathcal{E}}_{+,m}$, unique up to the addition of cusp forms, mapping to the canonical basis $\delta_{(r,n)}$ of $\mathbb{C}^{\mathcal{N}^0_{+,m}}$. Clearly, the space $\mathcal{E}^{\mathrm{OG,III}}_{+,m}/\mathcal{E}^0_{+,m}$ is spanned by $\psi_{m;(r,n)}$ with $0 < r < m$ and $4mn - r^2 = -1$, and has dimension 2^{s-1}. (There are exactly two square roots of 1 modulo $2P$ for each exact prime power divisor P of m and these can be combined arbitrarily by the Chinese remainder theorem. This gives 2^s residue classes $r \pmod{2m}$ with $r^2 \equiv 1 \pmod{4m}$, and if we choose the representatives of smallest absolute values then half of them are positive.) Moreover, the basis elements can be chosen to be a single orbit of the group $\{W_{m_1}\}_{m_1|m}$ introduced in §4.4, whose cardinality is the same number 2^{s-1}. Namely, if A_m is any choice of $\psi_{m;(1,0)}$, i.e. any function in $\widetilde{\mathcal{E}}_{+,m}$ whose unique non-zero coefficient in $\mathcal{N}^0_{+.m}$ is $c(0,1) = 1$, then the other functions $\psi_{m;(r,n)}$ with $(r,n) \in \mathcal{N}^0_{+,m}$ and $4mn - r^2 = -1$, can be chosen to be the images of A_m under the Atkin-Lehner involutions W_{m_1}, $m = m_1 m_2$. More explicitly, with this choice $\psi_{m;(r,n)}$ equals $A_m|W_{m_1}$ with $m_1 = \gcd\left(\frac{r+1}{2}, m\right)$, $m_2 = \gcd\left(\frac{r-1}{2}, m\right)$, and conversely $A_m|W_{m_1}$ for any decomposition $m = m_1 m_2$ equals $\psi_{m;(|r_0|,(r_0^2-1)/4m)}$ where r_0 ($= 1^*$ in the notation of §4.4) is the unique r_0 satisfying

$$r_0 \equiv -1 \pmod{2m_1}\ , \qquad r_0 \equiv 1 \pmod{2m_2}\ , \qquad |r_0| < m\ . \tag{10.29}$$

On the other hand, for any $\varphi \in \mathcal{E}^{\mathrm{OG,III}}_{+,m}$, we have

$$R_{1,m}(\varphi) \;=\; \mathcal{R}[\varphi] \;=\; 2c_\varphi(0,1) \tag{10.30}$$

(because $c_\varphi(0,-1) = c_\varphi(0,1)$ and $c_\varphi(0,r) = 0$ for all $r \neq \pm 1$ by the defining property of $\mathcal{E}^{\mathrm{OG,III}}_{+,m}$), and more generally,

$$R_{m_1,m_2}(\varphi) \;=\; 2\,c_\varphi(r_0,(r_0^2-1)/4m) \tag{10.31}$$

for each decomposition $m = m_1 m_2$, and with r_0 as in (10.29). It follows that, for any two decompositions $m = m_1 m_2$ and $m = m_1' m_2'$, we have

$$R_{m_1,m_2}\big(A_m\,|\,W_{m_1'}\big) \;=\; \begin{cases} 2 & \text{if } m_1' = m_1 \text{ or } m_1' = m_2\ , \\ 0 & \text{otherwise}\ . \end{cases} \tag{10.32}$$

This shows that the map

$$\mathcal{E}^{\mathrm{OG,III}}_{+,m}/\mathcal{E}^0_{+,m} \xrightarrow{R} \bigoplus_{\substack{m=m_1\cdot m_2 \\ \text{up to order}}} \mathbb{C} \tag{10.33}$$

is an isomorphism, and in combination with Theorem 10.4, completes the proof of the theorem in the case of square-free m and $\star = \mathrm{III}$: any $\varphi \in \widetilde{\mathcal{E}}_{+,m}$ can be decomposed, uniquely up to Jacobi cusp forms, as the sum of $\frac{1}{2}\sum_{m=m_1m_2} R_{m_1,m_2}(\varphi)A_m|W_{m_1} \in \mathcal{E}^{\mathrm{OG,III}}_{+,m}$ and an element of $\widetilde{J}_{2,m}$.

The proofs for $\star = \mathrm{I}$ and $\star = \mathrm{II}$ are almost identical, changing the basis $\{\psi_{m;(r,n)} = A_m|W_{m_1}\}$ of $\mathcal{E}^{\mathrm{OG,III}}_{+,m}/\mathcal{E}^{0}_{+,m}$ in both cases in such a way as to satisfy the new defining condition without changing the residues, i.e., in the case $\star = \mathrm{II}$ by replacing the first basis element $\psi_{m;(1,0)}$ by $2\psi_{m;(0,0)}$, and in the case $\star = \mathrm{I}$ by adding $2\psi_{m;(0,0)} - 2^{-s+1}\sum_{r'}\psi_{m;(r',n)}$ to each basis element.

The proof of the statement for forms of arbitrary index m is very similar, since there are still 2^{s-1} polar coefficients of a form in $\mathcal{E}^{\mathrm{OG}}_{+,m}$, in one-to-one correspondence with the decompositions of m into two coprime factors. The details are left to the reader. ☐

Theorem 10.5 says that the first line in the diagram

$$\begin{tikzcd}
0 \arrow[r] & J^{\mathrm{prim}}_{2,m} \arrow[r, "\mathrm{diag}"] \arrow[d, phantom, "\cap"] & \widetilde{J}^{\mathrm{prim}}_{2,m} \oplus \mathcal{E}^{\mathrm{OG,prim},\star}_{m} \arrow[r, "-"] \arrow[d, phantom, "\cap"] & \widetilde{\mathcal{E}}^{\mathrm{prim}}_{+,m} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & J^{\mathrm{OG,prim}}_{2,m} \arrow[r, "\mathrm{diag}"] & \widetilde{J}^{\mathrm{prim}}_{2,m} \oplus \mathcal{E}^{\mathrm{OG,prim}}_{m} \arrow[r, "-"] & \widetilde{\mathcal{E}}^{\mathrm{prim}}_{+,m} \arrow[r] & 0
\end{tikzcd} \tag{10.34}$$

is exact. The exactness of the second line is a trivial consequence, and the fact that $\mathcal{E}^{\mathrm{OG,prim},\star}_{+,m}$ has codimension exactly one in $\mathcal{E}^{\mathrm{OG,prim}}_{+,m}$ implies that $J^{\mathrm{prim}}_{2,m}$ has codimension one in $J^{\mathrm{OG,prim}}_{2,m}$. Moreover, the proof of the theorem also shows that $J^{\mathrm{OG},\star}_{2,m} = J_{2,m}$. (If $\varphi \in J^{\mathrm{OG}}_{2,m}$, then the vanishing of the residues $R_{m_1m_2}(\varphi)$ for all decompositions with $(m_1, m_2) = 1$ implies that the polar coefficients $C_\varphi(-1, r)$ are all equal to one another and to $-\frac{1}{2}c_\varphi(0,0)$, so that imposing any of the conditions (I), (II), (III) forces φ to be holomorphic. Conversely, if $\varphi \in J_{2,m}$, then the conditions OG, (I), (II) are trivial, and (III) holds because $\mathcal{R}[\varphi] = 0$.) It follows that the codimension of $J_{2,m}$ in $J^{\mathrm{OG}}_{2,m}$ is also one.

10.3 The residues of mock Jacobi forms

We can apply the general theorems of the previous subsection to mock Jacobi forms. In particular, if we have any weak mock Jacobi form of weight $k \geq 3$ and index m, it will be in $\widetilde{\mathcal{E}}_{\pm,m}$, and can be corrected by a weak Jacobi form to get a holomorphic mock Jacobi form. In the case $k = 2$, there is an obstruction $R(\varphi)$ to this being true, but, by the way we defined it, this obstruction can only depend on the shadow of φ. In this subsection, we show how to compute $R(\varphi)$ explicitly in terms of the shadow. Since $R(\varphi)$ is the vector with components $\mathcal{R}\big[\varphi|u_{(m_1,m_2)}|W_{m_1/(m_1,m_2)}\big]$, it suffices to compute $R[\varphi]$ for a weak mock Jacobi form of weight 2 in terms of the shadow of φ.

Let $\varphi(\tau, z)$ be a (strong or weak) mock Jacobi form of weight 2 and index m. Then, by definition, φ has a theta expansion (4.10) where each h_ℓ is a (in general weakly holomorphic) mock modular form of weight 3/2 and has a completion $\widehat{h}_\ell = h_\ell + g^*_\ell$ for some modular form g_ℓ of weight 1/2 such that the completion $\widehat{\varphi} = \sum_\ell \widehat{h}_\ell\,\vartheta_{m,\ell} = \varphi + \sum_\ell g^*_\ell\,\vartheta_{m,\ell}$ transforms like a Jacobi form of weight 2 and index m.

As usual, we denote by $\vartheta^0_{m,\ell}(\tau)$ the Thetanullwert $\vartheta_{m,\ell}(\tau,0)$.

Theorem 10.6. *Let φ be as above. Then*

$$\mathcal{R}[\varphi] \;=\; \frac{\sqrt{\pi}}{6}\sum_{\ell \;(\mathrm{mod}\; 2m)} \bigl(\vartheta^0_{m,\,\ell}, g_\ell\bigr)\,, \tag{10.35}$$

where $(\,\cdot\,,\,\cdot\,)$ denotes the Petersson scalar product in weight 1/2.

Proof. With the normalization of the Petersson scalar product as given in (3.14), we have

$$\sum_{\ell \;(\mathrm{mod}\; 2m)} \bigl(\vartheta^0_{m,\ell}, g_\ell\bigr) \;=\; \frac{3}{\pi}\int_{\mathcal{F}} F(\tau)\, d\mu(\tau)\;, \tag{10.36}$$

where

$$F(\tau) \;=\; \tau_2^{1/2}\sum_{\ell \;(\mathrm{mod}\; 2m)} \overline{g_\ell(\tau)}\,\vartheta^0_{m,\ell}(\tau)\;, \tag{10.37}$$

and $\mathcal{F}$ denotes a fundamental domain for $\Gamma \equiv SL_2(\mathbb{Z})$. This makes sense since the function $F(\tau)$ is Γ-invariant (the individual terms in (10.37) are not, which is why we had to be a little careful in the normalization of the scalar product) and is convergent because the weight of the modular forms g_ℓ and $\vartheta^0_{m,\ell}$ is less than 1. From eq. (7.1) it follows that the $\overline{\tau}$-derivative of the completion

$$\widehat{\varphi}(\tau,0) \;=\; \varphi(\tau,0) \,+\, \sum_{\ell \;(\mathrm{mod}\; 2m)} g^*_\ell(\tau)\,\vartheta^0_{m,\ell}(\tau)\,, \tag{10.38}$$

of $\varphi(\tau,0)$ satisfies

$$\frac{\partial}{\partial\overline{\tau}}\,\widehat{\varphi}(\tau,0) \;=\; \frac{1}{4i\sqrt{\pi}}\,\frac{F(\tau)}{\tau_2^2}\,. \tag{10.39}$$

The fact that $\widehat{\varphi}(\tau,0)$ transforms like a modular form of weight 2 on Γ implies that the differential one-form $\omega = \widehat{\varphi}(\tau,0)\,d\tau$ is Γ invariant, and equation (10.39) implies

$$d\omega \;=\; d\bigl(\widehat{\varphi}(\tau,0)\,d\tau\bigr) \;=\; -\frac{\partial\,\widehat{\varphi}(\tau,0)}{\partial\overline{\tau}}\,d\tau\,d\overline{\tau} \;=\; \frac{1}{2\sqrt{\pi}}\,F(\tau)\,d\mu(\tau)\;. \tag{10.40}$$

Therefore, by Stokes's theorem, we have

$$\frac{\sqrt{\pi}}{6}\sum_{\ell \;(\mathrm{mod}\; 2m)} \bigl(\vartheta^0_{m,\ell}, g_\ell\bigr) \;=\; \int_{\mathcal{F}} d\omega \;=\; \int_{\partial\mathcal{F}} \omega \;=\; \lim_{T\to\infty}\int_{\partial\mathcal{F}(T)} \omega \tag{10.41}$$

where $\mathcal{F}(T)$ is the truncated fundamental domain

$$\{\tau \in \mathcal{H} \mid |\tau| > 1\,,\; |\tau_1| < \frac{1}{2}\,,\; \tau_2 < T\}\;. \tag{10.42}$$

On the other hand, we have

$$\int_{\partial\mathcal{F}(T)}\omega \;=\; \int_{-\frac12+iT}^{\frac12+iT}\omega \;=\; \int_{-\frac12+iT}^{\frac12+iT}\varphi(\tau,0)\,d\tau \;+\; O\Bigl(\frac{1}{\sqrt{T}}\Bigr)\,. \qquad (10.43)$$

The first equality holds because ω is Γ-invariant, and all the edges of $\mathcal{F}(T)$ except the top edge $\tau_2 = T$ come in pairs which are Γ-equivalent, but of opposite orientation. The second equality follows from (10.38), and because $\vartheta^0_{m,\ell}$ is $O(1)$ and $g^*_\ell(\tau)$ is $O(T^{-1/2})$ by (7.2) with $k = 3/2$. The theorem then follows. ☐

The theorem immediately implies a more general formula for the residue $R_{m_1,m_2}(\varphi) = \mathcal{R}\bigl[\varphi\big|W_{m_1}\bigr]$ when $m = m_1m_2$ with $(m_1, m_2) = 1$ in terms of Petersson scalar products, namely,

$$R_{m_1,m_2}(\varphi) \;=\; \frac{\sqrt{\pi}}{6}\sum_{\ell\;(\mathrm{mod}\;2m)}\bigl(\vartheta^0_{m,\,\ell},\, g_{\ell^*}\bigr)\,, \qquad (10.44)$$

where $\ell \mapsto \ell^*$ is the involution on $(\mathbb{Z}/2m\mathbb{Z})^*$ appearing in the definition of W_{m_1}.

10.4 Remaining proofs of theorems from §9

In this subsection, we apply the results proved so far to the special family studied in §9, namely, the meromorphic Jacobi forms $\varphi_{2,m}$ having a pole $1/(2\pi iz)^2$ at the origin and their associated mock Jacobi forms $\Phi_{2,m}$. Specifically, we shall prove Theorems 9.1, 9.2, 9.4, 9.5, and 9.7.

We begin with a preliminary remark. Two of the results we need to prove are Theorem 9.1 and the statement that the functions $\Phi^0_{2,m}$ defined implicitly by (9.37) are primitive modulo weak Jacobi forms. Both of these have the form that a certain elliptic form is in fact a (weak) Jacobi form, i.e., that its shadow vanishes. Since the forms in question are defined using the Hecke-like operators defined in §4.4, we first have to explain how the action of these operators extends to non-holomorphic elliptic forms, since then we can apply them to the completions of our mock Jacobi forms and verify the necessary identities between shadows. The definitions of the Hecke-like operators were given in terms of the Fourier coefficients $c(n, r)$ or $C(\Delta, \ell)$ defined by (4.3) and (4.4). In the case of the completions of mock modular forms Φ, when we apply equation (4.36) or (4.37) for the action of U_t or V_t with the φ there replaced by $\Phi^C := \widehat{\Phi} - \Phi$ (the "correction term" discussed at the end of §8.3), we must interpret $c(n, r)$ as the coefficient of $\beta(|n|\tau_2)\,q^n\,y^r$ rather than simply q^ny^r as in the holomorphic case, where $\beta(t) = \sqrt{\pi/4m}\;\mathrm{erfc}(2\sqrt{\pi t})$ in the weight 2 case, and similarly for equation (4.43) and $C(\Delta,\, r\;(\mathrm{mod}\;2m))$. In view of (9.5) and (7.2), these coefficients in the case of $\Phi_{2,m}$ are given by

$$C(\Phi^C_{2,m};\Delta,\,\ell\;(\mathrm{mod}\;2m)) \;=\; \sum_{\substack{\lambda^2=-\Delta\\ \lambda\equiv\ell\;(\mathrm{mod}\;2m)}}|\lambda|\,. \qquad (10.45)$$

Proof of Theorem 9.1: We have to show that for $m = P_1 \cdots P_r$ ($P_i = p_i^{\nu_i}$, p_i distinct primes) we have

$$\Phi_{2,m} \,|\, \prod_{i=1}^{r}(1+\varepsilon_i W_{P_i}) \;\equiv\; \Bigl(\Phi_{2,M} \,|\, \prod_{p|M}(1-W_p)\Bigr) \,|\, \mathcal{V}^{(M)}_{2,m/M}\,, \tag{10.46}$$

where M is the product of the primes p_i for which $\varepsilon_i = -1$ and "$\equiv$" means that the two sides of the equation differ by a weak Jacobi form. Because of the multiplicativity of the Hecke-like operators U_d, V_t and W_{m_1}, this formula can be rewritten as

$$\Phi_{2,m} \,|\, \prod_{i=1}^{r}(1+\varepsilon_i W_{P_i}) \;\equiv\; \Phi_{2,M} \,|\, \prod_{\varepsilon_i=-1}(1-W_{p_i})\mathcal{V}^{(p_i)}_{P_i/p_i} \,|\, \prod_{\varepsilon_i=+1} \mathcal{V}^{(1)}_{P_i}\,, \tag{10.47}$$

so by induction on r, it suffices to show that if $P = p^\nu || m$ then

$$\Phi_{2,m} \,|\, (1\,+\,W_P) \;\equiv\; \Phi_{2,m/P} \,|\, \mathcal{V}^{(1)}_{2,P}\,, \qquad \Phi_{2,m} \,|\, (1\,-\,W_P) \;\equiv\; \Phi_{2,mp/P} \,|\, (1-W_p)\,\mathcal{V}^{(p)}_{2,P/p}\,. \tag{10.48}$$

For this, it is enough to show that in each the difference of the left-hand side and the right-hand side has shadow zero.

For the first equation in (10.48), eq. (10.45) implies that the "correction term" for the left-hand side is given by

$$\begin{aligned} &C(\Phi^C_{2,m}|(1+W_P)\,;\Delta,\ell \text{ (mod } 2m)) = C(\Phi^C_{2,m}\,;\Delta,\ell \text{ (mod } 2m)) + C(\Phi^C_{2,m}\,;\Delta,\ell^* \text{ (mod } 2m)) \\ &= \begin{cases} D(\delta^{(2m)}_{D,\ell} + \delta^{(2m)}_{D,-\ell} + \delta^{(2m)}_{D,\ell^*} + \delta^{(2m)}_{D,-\ell^*}) & \text{if } \Delta = -D^2, \text{ for some } D \in \mathbb{N} \\ 0 & \text{if } -\Delta \text{ is not a perfect square} \end{cases} \end{aligned} \tag{10.49}$$

where ℓ^* (mod $2m$) is defined by $\ell^* \equiv -\ell$ (mod $2P$), $\ell^* \equiv +\ell$ (mod $2m/P$), and the notation $\delta^{(n)}_{a,b}$ means $\delta_{a,b \text{ (mod } n)}$. The correction term for the right-hand side, on the other hand, is given (since $\mathcal{V}^{(1)}_{2,P} = V_{2,P} - pV_{2,P/p^2}U_p$, where the second term is omitted if $p^2 \nmid P$) by

$$\begin{aligned} &\sum_{d|(\frac{\Delta+\ell^2}{4m},\ell,P)} d\,C\Bigl(\Phi^C_{2,\frac{m}{P}}\,;\frac{\Delta}{d^2},\frac{\ell}{d}\text{ (mod } 2\frac{m}{P})\Bigr) \quad - p \sum_{d|(\frac{\Delta+\ell^2}{4m},\ell,\frac{P}{p^2})} d\,C\Bigl(\Phi^C_{2,\frac{m}{P}}\,;\frac{\Delta}{p^2d^2},\frac{\ell}{pd}\text{ (mod } 2\frac{m}{P})\Bigr) \\ &\quad = \sum_{d|(\frac{\Delta+\ell^2}{4m},\ell,P)} \;\sum_{\substack{\lambda^2=-\Delta \\ \lambda\equiv\ell \text{ (mod } \frac{2md}{P})}} |\lambda| \quad - \sum_{d|(\frac{\Delta+\ell^2}{4m},\ell,\frac{P}{p^2})} \;\sum_{\substack{\lambda^2=-\Delta \\ \lambda\equiv\ell \text{ (mod } \frac{2md}{P})}} |\lambda|\,, \\ &\quad = \sum_{\substack{d\in\{P,P/p\} \\ d|(\frac{\Delta+\ell^2}{4m},\ell,P)}} \;\sum_{\substack{\lambda^2=-\Delta \\ \lambda\equiv\ell \text{ (mod } \frac{2md}{P})}} |\lambda|\,, \end{aligned} \tag{10.50}$$

where the last line holds because the terms in the first sum in the second line with $d|Pp^{-2}$ cancel with the terms in the second sum. The terms with $d = P$ in the third line give the first two terms $\sqrt{-D} = \pm\ell^*$ of the right hand side of (10.49), while the terms with $d = P$ in the third line give the second two terms $\sqrt{-D} = \pm\ell^*$.

The correction term for the left-hand side of the second equation in (10.48) is given by the the right-hand side of (10.49) with the signs of the last two terms changed. This is now equal to zero unless $(\ell, m) = 1$, in which case it is equal to $\sqrt{-\Delta}$ (resp. $-\sqrt{-\Delta}$) if $D \equiv \pm\ell \pmod{2m}$ (resp. $D \equiv \pm\ell^* \pmod{2m}$). The correction term for the right-hand side, on the other hand, is given by

$$\begin{aligned}\sum_{d|(\ell,\frac{P}{p},\frac{\Delta+\ell^2}{4m})} d\,\biggl(C\Bigl(\Phi^C_{2,mp/P}\,;\frac{\Delta}{d^2}\,,\frac{\ell}{d}\ \bigl(\mathrm{mod}\ \frac{2mp}{P}\bigr)\Bigr)\;-\;C\Bigl(\Phi^C_{2,mp/P}\,;\frac{\Delta}{d^2}\,,\frac{\ell^*}{d}\ \bigl(\mathrm{mod}\ \frac{2mp}{P}\bigr)\Bigr)\biggr) \\ =\ \sum_{d|(\ell,\frac{P}{p},\frac{\Delta+\ell^2}{4m})}\Biggl(\ \sum_{\substack{\lambda^2=-\Delta\\ \frac{\lambda}{d}\equiv\frac{\ell}{d}\ (\mathrm{mod}\ \frac{2mp}{P})}}\ -\ \sum_{\substack{\lambda^2=-\Delta\\ \frac{\lambda}{d}\equiv\frac{\ell^*}{d}\ (\mathrm{mod}\ \frac{2mp}{P})}}\ \Biggr)|\lambda|\,,\end{aligned}\tag{10.51}$$

where $\ell^* \pmod{2mp/P}$ is now defined by $\ell^* \equiv -\ell \pmod{2p}$, $\ell^* \equiv +\ell \pmod{2m/P}$. If $p|\ell$, then $\ell^* = \ell$ and the expression vanishes, so that only $d = 1$ contributes in the first summation. The condition $4m|(\Delta + \ell^2)$ combined with $\lambda \equiv \ell \pmod{2mp/P}$ implies $\lambda \equiv \ell \pmod{2m}$ (and similarly for ℓ^* since $\ell^2 \equiv \ell^{*2} \pmod{2mp/P}$), thus proving the second equation in (10.48). $\square$

Proof of Theorem 9.2: By Möbius inversion, the mock Jacobi forms $\Phi_{2,m}$ associated to any choice of minimal meromorphic Jacobi forms $\varphi_{2,m}$ can be represented in the form (9.37), where $\Phi^0_{2,m}$ is defined by

$$\Phi^0_{2,m} \;=\; \sum_{d^2|m} \mu(d)\ \Phi_{2,m/d^2}\,|\,U_d\,. \tag{10.52}$$

We have to show that $\Phi^0_{2,m}$ can be chosen to be primitive and to have optimal growth.

To prove the second of these statements, it is enough to show that $\Phi^0_{2,m}|u_p$ has no shadow for all primes p with $p^2|m$. Then, by Theorem 10.5, the weak elliptic form $\Phi^0_{2,m}$ can be chosen (by adding a weak Jacobi form) to have optimal growth. Writing the square divisors of m as d^2p^{2i} with $p \nmid d$ and noting that $\mu(dp^i) = 0$ for $i > 1$ and that $U_{dp} = U_dU_p$, $U_du_p = u_pU_d$, and $U_pu_p = p$, we find from (10.52) that

$$\begin{aligned}\Phi^0_{2,m}\,|\,u_p \;&=\; \sum_{d^2|m,\,p\nmid d} \mu(d)\bigl[\Phi_{2,m/d^2}\,|\,U_d \;-\; \Phi_{2,m/d^2p^2}\,|\,U_d\,|\,U_p\bigr]\;|\;u_p \\ &=\; \sum_{d^2|m,\,p\nmid d} \mu(d)\bigl[\Phi_{2,m/d^2}\,|\,u_p \;-\; p\,\Phi_{2,m/d^2p^2}\bigr]\;|\;U_d\,,\end{aligned}$$

so it suffices to show that $\Phi_{2,m}|u_p$ and $p\,\Phi_{2,m/p^2}$ have the same shadow for any m with $p^2|m$. But the shadow of $\Phi_{2,m} = \varphi^{\mathrm{F}}_{2,m}$ is determined by $\varphi^C_{2,m} = \widehat{\Phi}_{2,m} - \Phi_{2,m}$ (cf. comments at the end of §8.3), and by (10.45), we have:

$$C(\varphi^C_{2,m}\,;\Delta p^2,\,rp\ (\mathrm{mod}\ 2m))\ =\ \begin{cases} p|\lambda| & \text{if } \Delta=-\lambda^2,\quad \lambda\equiv r\ (\mathrm{mod}\ 2m/p)\,,\\ 0 & \text{otherwise}\,.\end{cases} \tag{10.53}$$

Then, from (4.43), we have:

$$\begin{aligned} C(\varphi^C_{2,m}|u_p\,;\Delta,\,\ell\ (\mathrm{mod}\ 2m/p^2)) \ &=\ \sum_{\substack{r\ (\mathrm{mod}\ 2m/p)\\ r\equiv\ell\ (\mathrm{mod}\ 2m/p^2)}} C(\varphi^C_{2,m}\,;\Delta p^2,\,rp\ (\mathrm{mod}\ 2m))\,,\\ &=\ \begin{cases} p|\lambda| & \text{if } \Delta=-\lambda^2,\quad \lambda\equiv\ell\ (\mathrm{mod}\ 2m/p^2)\,,\\ 0 & \text{otherwise}\,.\end{cases}\\ &=\ p\,C(\varphi^C_{2,m/p^2}\,;\Delta,\ell)\,. \end{aligned} \tag{10.54}$$

This completes the proof of the fact that $\Phi^0_{2,m}$ can be chosen to have optimal growth. Now, by formula (4.43), we have that $\Phi^0_{2,m}|u_t$, and hence also $\Phi^0_{2,m}|u_t|U_t$, is holomorphic for $t>1$, so by replacing $\Phi^0_{2,m}$ by $\pi^{\mathrm{prim}}(\Phi^0_{2,m})$ as defined in (4.46) we can also assume that $\Phi^0_{2,m}$ is primitive. ☐

Proof of Theorem 9.4: The proof of Theorem 10.5 given in §10.2 the following discussion already showed that $J^{\mathrm{OG}}_{2,m}/J_{2,m}$ is one-dimensional and that any representative $\mathcal{K}_m$ of this quotient space, if we normalize it to have $c(\mathcal{K}_m\,;0,0)=2$, has polar coefficients as given in the statement of Theorem 9.4. Since all the polar coefficients $C(\mathcal{K}_m\,;-1,r)$ are equal, we have that $\mathcal{K}_m|W_{m_1}-\mathcal{K}_m\in J_{2,m}$ for all $m_1||m$, so (by averaging) we can choose $\mathcal{K}_m$ to be invariant under all Atkin-Lehner operators. Finally, we can make $\mathcal{K}_m$ primitive by the argument used at the end of the proof of Theorem 9.2. ☐

Remark. For square-free m, since all holomorphic Jacobi forms are in fact cusp forms, $\mathcal{K}_m$ is unique up to Jacobi cusp forms, and its coefficients with $\Delta=0$ are unique. In this case the only such coefficient, up to translation, is $c(0,0)$. If m is not square-free, there is more than one coefficient $c(n,r)$ such that $\Delta=0$. Notice that $r^2\equiv 0\ (\mathrm{mod}\ 4m)\Leftrightarrow(r,2m)=2m/t$ with $t^2|m$. For each t with $t^2|m$, we have

$$\sum_{j\ (\mathrm{mod}\ t)} C(\mathcal{K}_m\,;0,\frac{2mj}{t}\ (\mathrm{mod}\ 2m))\ =\ \begin{cases} -2 & \text{if } t=1\,,\\ 0 & \text{if } t>1\,.\end{cases} \tag{10.55}$$

(For $t=1$, the sum on the left reduces to the single term $c(\mathcal{K}_m\,;0,0)=-2$, and for $t>1$, it equals $C(\mathcal{K}_m|u_t;0,0)$, which vanishes because $\mathcal{K}_m|u_t\in J_{2,m/t^2}$.) We can choose the $\Delta=0$ coefficients arbitrarily subject to this one constraint, and once we have made a choice, then $\mathcal{K}_m$ is unique up to cusp forms even in the non-square-free case. For instance, at the expense of introducing denominators, we could fix $\mathcal{K}_m$ uniquely up to cusp forms by requiring that $C(\mathcal{K}_m;0,r)$

depends only on the gcd of r and $2m$. (This would correspond to choosing $a = \frac{1}{2}$ in the first and third tables at the end of §9.3.) Another choice, more similar to our "$\star = \mathrm{II}$" condition, would be to choose $C(\mathcal{K}_m; 0, r) = 0$ for all r with $4m|r^2$ and $r \nmid 2m$, in which case the remaining coefficients $C(\mathcal{K}_m; 0, 2m/t)$ $(t^2|m)$ are uniquely determined by (10.55). (This would correspond to choosing $a = 1$ in the first of the tables at the end of §9.3.) The reader can easily work out the complete formula for all of the coefficients $C(\mathcal{K}_m; 0, r)$ for either of these two special choices.

Our next task is to prove Theorem 9.5. For this we need to calculate the residues of $\Phi_{2,m}$, using the results of the previous subsection. We therefore have to apply the formula (10.44) to the mock modular form $\Phi_{2,m} = \varphi^{\mathrm{F}}_{2,m}$. (Note that the residue of this form will be independent of the specific choice of $\varphi_{2,m}$, since any two choices differ by a Jacobi form and the residues of Jacobi forms vanish.) Since the completion of $\Phi_{2,m}$ involves the weight $1/2$ unary theta series $\vartheta^0_{m,\ell}$, the first step is to compute the scalar products of these theta series.

Proposition 10.2. *For $m \in \mathbb{N}$, and $\ell_1, \ell_2 \in \mathbb{Z}/2m\mathbb{Z}$, we have*

$$(\vartheta^0_{m,\ell_1}\,,\,\vartheta^0_{m,\ell_2}) \;=\; \frac{1}{2\sqrt{m}}\big(\delta_{\ell_1,\ell_2\,(\mathrm{mod}\ 2m)} + \delta_{\ell_1,-\ell_2\,(\mathrm{mod}\ 2m)}\big) \tag{10.56}$$

Proof. Formula (3.15) gives

$$(\vartheta^0_{m,\ell_1}\,,\,\vartheta^0_{m,\ell_2}) \;=\; \frac{1}{2}\,\mathrm{Res}_{s=\frac{1}{2}}\left(\sum_{r>0}\frac{(\delta_{r,\ell_1\,(2m)} + \delta_{r,-\ell_1\,(2m)})(\delta_{r,\ell_2\,(2m)} + \delta_{r,-\ell_2\,(2m)})}{(r^2/4m)^s}\right).$$

The propostition follows. ☐

Proposition 10.3. *Let $\varphi_{2,m}$ be any meromorphic Jacobi form of weight 2 and index m with pole $1/(2\pi iz)^2$ and its translates, and let $\Phi_{2,m}$ be its finite part $\varphi^{\mathrm{F}}_{2,m}$. Then*

$$R_{m_1,m_2}(\Phi_{2,m}) \;=\; \frac{m_1+m_2}{12}\,. \tag{10.57}$$

Proof. By equation (9.5), the ℓ^{th} component g_ℓ of the shadow for $\Phi_{2,m}$ is $-\sqrt{\frac{m}{4\pi}}\vartheta^0_{m,\ell}$. When $(m_1,m_2) = 1$, $R_{m_1,m_2}(\Phi_{2,m}) = \mathcal{R}[\Phi_{2,m}\,|\,W_{m_1}]$. Theorem 10.6 then gives:

$$\begin{aligned}\mathcal{R}[\Phi_{2,m}\,|\,W_{m_1}] \;&=\; \frac{\sqrt{m}}{\sqrt{4\pi}}\frac{\sqrt{\pi}}{6}\sum_{\ell\,(\mathrm{mod}\ 2m)}\big(\vartheta^0_{m,\ell},\vartheta^0_{m,\ell^*}\big)\\ &=\; \frac{1}{24}\sum_{\ell\,(\mathrm{mod}\ 2m)}\big(\delta_{\ell,\ell^*\,(\mathrm{mod}\ 2m)} + \delta_{\ell,-\ell^*\,(\mathrm{mod}\ 2m)}\big)\\ &=\; \frac{1}{12}\,(m_1+m_2)\,.\end{aligned}$$

In this calculation, we used Proposition 10.2 to get the second line. For the third line, recall that ℓ^* is defined by $(\ell^* \equiv -\ell \pmod{2m_1}, \ell^* \equiv \ell \pmod{2m_2})$, so that $\ell^* \equiv \ell \pmod{2m} \Leftrightarrow m_1\,|\,\ell$ (which is true for $2m_2$ values of ℓ in $\mathbb{Z}/2m\mathbb{Z}$), and similarly, $\ell^* \equiv -\ell \pmod{2m} \Leftrightarrow m_2\,|\,\ell$ (which is true for $2m_1$ values of ℓ). When $(m_1, m_2) = t > 1$, we have $R_{m_1,m_2}(\Phi_{2,m}) = R_{m_1/t,m_2/t}(\Phi_{2,m}\,|\,u_t)$. As we have shown in the proof of Theorem 9.2, the shadow of $\Phi_{2,m}\,|\,u_t$ is the same as the shadow of $t\Phi_{2,m/t^2}$, so $R_{m_1/t,m_2/t}(\Phi_{2,m}\,|\,u_t) = tR_{m_1/t,m_2/t}(\Phi_{2,m/t^2}) = t(\frac{m_1}{t} + \frac{m_2}{t})/12 = (m_1+m_2)/12$ by the special case already proved. ☐

The reader might find it helpful to compare the statement of this proposition with the table in §9.2 presenting the data for the case $m = 6$, in which the function $\Phi = \Phi_{2,6}$ (in any of its three versions I, II, or III) has

$$\begin{aligned} R_{1,6}(\Phi) \;&=\; 2C(\Phi; -1, \pm 1 \,(\mathrm{mod}\ 12)) + C(\Phi; 0, 0\,(\mathrm{mod}\ 12)) \;=\; 7/12\,, \\ R_{2,3}(\Phi) \;&=\; 2C(\Phi; -1, \pm 5 \,(\mathrm{mod}\ 12)) + C(\Phi; 0, 0\,(\mathrm{mod}\ 12)) \;=\; 5/12\,. \end{aligned}$$

Proof of Theorem 9.5: We can now complete the proof of Theorem 9.5 and its corollary. For m square-free, any choice of $\Phi = \Phi^0_{2,m}$ in $\mathcal{E}^{\mathrm{OG}}_m$, and any r with $r^2 \equiv 1 \pmod{4m}$, we have

$$2C(\Phi;\,-1,\ell) + C(\Phi;\,0,0) \;=\; \mathcal{R}[\Phi\,|\,W_{m_1}] \;=\; \frac{m_1+m_2}{12} \tag{10.58}$$

by Proposition 10.3, where $m = m_1 m_2$ is the decomposition of m for which $r^* = 1$. Since $C(\Phi^{\mathrm{III}}; 0, 0) = 0$, this proves Theorem 9.5 and also gives the corresponding formulas

$$C(\Phi^{\mathrm{I}}; -1, r) \;=\; -\frac{(m_1-1)(m_2-1)}{12}\,, \qquad C(\Phi^{\mathrm{II}}; -1, r) \;=\; -\frac{1}{12}\left(\frac{m_1+m_2}{2} - \prod_{p|m}\frac{p+1}{2}\right)$$

for Φ^{I} and Φ^{II}, as well as formula (9.43) for the polar coefficients of $\mathcal{Q}_M$ when $\mu(M) = 1$. If m is not square-free, then the solutions $r \pmod{2m}$ of $r^2 \equiv 1 \pmod{4m}$ are still in 1:1 correspondence with the decompositions of m into coprime factors m_1 and m_2. To compute the corresponding Fourier coefficient of $\Phi^0_{2,m}$, we look at each term in (10.52) separately, writing each d as $d_1 d_2$ with $d_i^2 | m_i$ and noting that $\frac{m_1}{d_1^2} = (\frac{r-1}{2}, \frac{m}{d^2})$, $\frac{m_2}{d_2^2} = (\frac{r+1}{2}, \frac{m}{d^2})$, after which (9.41) follows from (10.57) and the fact that $\mathcal{R}[\varphi\,|\,U_d] \;=\; \mathcal{R}[\varphi]$. ☐

Proof of Theorem 9.7: It is enough to show that any weak Jacobi form of optimal growth is actually holomorphic since, as we have seen in Theorem 10.3, there are no holomorphic Jacobi forms of weight 1. To show this, we use the "residue operator on weight 1 forms" defined by

$$\mathcal{R}^{(1)}[\varphi] \;=\; \mathrm{Res}_{\tau=\infty}\big(\varphi'(\tau,0)\,d\tau\big) \;=\; \sum_{r\in\mathbb{Z}} r\,c_\varphi(0,r) \tag{10.59}$$

instead of (10.17), where, as usual, $\varphi'(t,z) = \frac{1}{2\pi i}\frac{d}{dz}\,\varphi(t,z)$. This residue vanishes if $\varphi \in \widetilde{J}_{1,m}$, since then $\varphi'(\cdot,0) \in M^!_2$. If, further, φ has optimal growth, then $2c_\varphi(0,1) = \mathcal{R}^{(1)}[\varphi] = 0$, so $C_\varphi(-1,1)$ vanishes. By the same argument with φ replaced by φ acted upon by each of the Atkin-Lehner operators, one has that $C_\varphi(-1,r) = 0$ for all r with $r^2 \equiv -1 \pmod{4m}$, and thus $\varphi \in J_{1,m}$. $\square$

11. Quantum black holes and mock modular forms

We now have all the ingredients required to address the questions posed in the introduction. These answers follow from applying our results of §8, in particular theorem 8.3, to the quarter-BPS partition function of the the $\mathcal{N} = 4$ theory. It turns out that in this case, many of the mathematical objects in the theorem simplify. A reader who has not gone through the details of §8 could still fruitfully read this section to obtain the relevant results for this example.

In particular, we show in §11.1 that the immortal black hole degeneracies extracted from the asymptotic partition function (6.14) are Fourier coefficients of a (mixed) mock Jacobi form. The completion transforms as a true Jacobi form and obeys a first order partial differential equation (11.23) which can be understood as a holomorphic anomaly equation. In §11.2 we discuss the physical origin of the meromorphy of the asymptotic counting function in this example and how it is related to the noncompactness of the microscopic brane dynamics. In §11.3 we analyze the contour corresponding to the M-theory limit to embed our analysis in the context of AdS_3/CFT_2 holography near the horizon of a black string. We conclude in §11.4 with comments and open problems.

11.1 Mock Jacobi forms, immortal black holes, and AdS_2 holography

We first discuss the consequences of the S-duality symmetry and its relation to the spectral flow symmetry for the moduli-independent immortal degeneracies $d^*(n,\ell,m)$ defined by (6.14). Recall that under the S-duality transformations

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \tag{11.1}$$

with integer b, the S-modulus (2.21) transforms as:

$$S \to S + b\,. \tag{11.2}$$

which means that the axion transforms as $a \to a + b$. Since the axion arises from a dimensional reduction of the 3-form field in the M-theory frame (discussed in §11.3), this subgroup (11.1) of S-duality has its origins in large gauge transformations of the 3-form field. When the axion transforms as above, the effective theta angle changes and as a result the charge vector of the dyon transforms as $(Q,P) \to (Q + Pb, P)$ due to the Witten effect [122]. Consequently, the charge-invariants transform as

$$n \to n + mb^2 + b\ell\,, \qquad \ell \to \ell + 2mb\,, \qquad m \to m. \tag{11.3}$$

which are nothing but the spectral flow transformations. Note that m is left fixed by this subgroup of S-duality whereas other elements of the S-duality group would change m. Thus, only this subgroup is expected to act on the sector with a fixed m. The covariance of the spectrum under these S-duality transformations implies

$$d(n,\ell,m)|_{S,\mu} \;=\; d(n+mb^2+b,\ell+2mb,m)|_{S+b,\mu}\,. \tag{11.4}$$

Since the degeneracies of immortal black holes are independent of the moduli by definition, they are actually invariant under S-duality:

$$d^*(n+mb^2+b,\ell+2mb,m) \;=\; d^*(n,\ell,m)\,. \tag{11.5}$$

From the mathematical point of view, this is precisely the action of the elliptic part of the Jacobi transformations (4.2) on the Fourier coefficients of a Jacobi (or mock Jacobi) form.

The immortal degeneracies $d^*(n,\ell,m)$ are computed using the attractor contour (6.13), for which the imaginary parts of the potentials are proportional to the charges and scale as

$$\mathrm{Im}(\sigma) \;=\; 2n/\varepsilon\,, \qquad \mathrm{Im}(\tau) \;=\; 2m/\varepsilon\,, \qquad \mathrm{Im}(z) \;=\; -\ell/\varepsilon\,, \tag{11.6}$$

with ε very small and positive. In other words,

$$|p| = \lambda^{2n}\,, \quad |q| = \lambda^{2m}\,, \quad |y| = \lambda^{-\ell}\,, \qquad \text{with } \lambda = \exp(-2\pi/\epsilon) \to 0 \tag{11.7}$$

on the attractor contour. We assume that $n > m$ without loss of generality, otherwise one can simply exchange σ and τ. Moreover, using the spectral flow symmetry (11.5) we can always bring ℓ to the window $0 \le \ell < 2m$ for the immortal degeneracies.

To extract the Fourier coefficient from the product representation (5.13) of the counting function (5.14) using (11.7), we need to expand a typical term in the product of the form

$$\frac{1}{(1-p^r q^s y^t)^{2C_0(4rs-t^2)}}\,. \tag{11.8}$$

where $2C_0$ is the Fourier coefficient of the elliptic genus of a single copy of $K3$. Depending on whether $|p^r q^s y^t|$ is less than one or greater than one, we expand the denominator using

$$\frac{1}{1-x} \;=\; \begin{cases} 1+x+x^2+\ldots & \text{for } |x|<1\,, \\ -(x^{-1}+x^{-2}+x^{-3}+\ldots) & \text{for } |x|>1\,. \end{cases} \tag{11.9}$$

We see that it is not *a priori* obvious that expanding first in p is the same as using the attractor contour because if the powers of t are sufficiently large and positive, the attractor contour may correspond to expanding some terms effectively in p^{-1} instead of in p.

To address this question we now show that for the terms that appear in the product representation, $|p^r q^s y^t|$ is always less than one using the fact that

$$C_0(\Delta) = 0 \qquad \text{for } \Delta < -1\,. \tag{11.10}$$

To begin with, the p-dependent terms with $\Delta = 0, -1$ arise only for $s = 0$ and are of the form

$$\frac{1}{(1-py)^2(1-py^{-1})^2(1-p^r)^{20}} \tag{11.11}$$

These involve only single powers y or y^{-1} and positive powers of p. The absolute value of py or py^{-1} for the attractor contour goes as $\lambda^{2m\pm l}$ which is always much less than one for $0 \leq |l| < 2m$. The remaining p-dependent terms with $s > 0$ have $\Delta = 4rs - t^2 > 0$ and $r > 0$. The absolute value of $p^r q^s y^t$ goes as

$$\lambda^{2nr+2ms-\ell t}\,. \tag{11.12}$$

The exponent is most negative when ℓ takes its maximum positive value $\ell_{max} = \sqrt{2mn}$ and t takes its maximum positive value $t_{max} = \sqrt{2rs}$. For black hole configurations with a smooth horizon, $4mn - \ell^2 > 0$ because the classical area of the horizon is the square-root of this quantity. Hence the most negative value of the exponent would be

$$2nr + 2ms - 2\sqrt{mnrs} \;=\; 2(\sqrt{nr} - \sqrt{ms})^2 > 0\,, \tag{11.13}$$

and the absolute value of $p^r q^s y^t$ is always less than or equal to

$$\lambda^{2(\sqrt{nr}-\sqrt{ms})^2} \tag{11.14}$$

which is always much less than one. This implies that to extract the attractor Fourier coefficients, we expand all terms in the product in the small p expansion using (11.9) thus giving us the Jacobi form ψ_m of (1.1) as a partition function in the remaining two chemical potentials.

The degeneracies extracted from $\psi_m(\tau, z)$ still experience many wall-crossings, and we should now use the rest of the attractor contour to analyze the immortal part of the degeneracies of ψ_m. We do this in two steps by first performing the inverse Fourier transform in z, and then in τ. For the first step, we need to specify a contour for the imaginary part of z relative to that of τ. The conditions (11.6) on the imaginary parts of (τ, z) implies that $\mathrm{Im}(z) = -\frac{\ell}{2m}\mathrm{Im}(\tau)$. Since by moving the contour in a purely horizontal direction we do not cross any poles, we can add an arbitrary real part to z, which we can choose so that $z = -\ell\tau/2m$. This gives the result of the first integral to be:

$$f^*_{m,\ell}(\tau) \;=\; e^{-\pi i l^2\tau/2m} \int_P^{P+1} \psi_m(\tau, z)\, e^{-2\pi i \ell z}\, dz\,, \qquad P = -\ell\tau/2m\,. \tag{11.15}$$

We then have to perform a τ integral with the contour specified by the attractor values (11.6), but since the integrand $f^*_{m,\ell}(\tau)$ is holomorphic in τ, the answer is independent of the contour. We thus get the degeneracies $d^*(n,\ell,m)$ to be:

$$d^*(n,\ell,m) \;=\; \int e^{\pi i\ell^2\tau/2m}\, f^*_{m,\ell}(\tau)\, d\tau\,, \tag{11.16}$$

where the integral is over an interval of length 1 for the fixed imaginary value (11.6) of τ.

To extend to other values of ℓ, we use the spectral flow invariance (11.5) and sum over all values of ℓ to get a canonical partition function in the two variables (τ, z). Because of the factor $e^{\pi i\ell^2\tau/2m}$ in (11.16), the sum over spectral-flow images yields the function $\vartheta_{m,l}(\tau, z)$. Putting all this together, and comparing the above equations (11.15), (11.16) with our definition (8.3) of the finite part of a meromorphic Jacobi form, we see that the single centered black hole degeneracies are Fourier coefficients of precisely the *finite* part of ψ_m as defined in §8:

$$\sum_{\ell \bmod (2m)} f^*_{m,\ell}(\tau)\, \vartheta_{m,\ell}(\tau,z) \;=\; \psi^{\mathrm{F}}_m(\tau,z)\,. \tag{11.17}$$

We see that the seemingly unconventional charge-dependent choice for the contour in §8 for defining the finite part is completely natural from the point of view of the attractor mechanism (indeed, it was motivated there by these physical considerations). According to our analysis in §8, this finite part ψ^{F}_m is a *mock* Jacobi form. This is one of the main conclusions of our analysis, so we summarize the entire discussion by the following concise statement:

> *The degeneracies of single centered (immortal) black holes with magnetic charge invariant $M^2/2 = m$ are Fourier coefficients of a mock Jacobi form of index m.*

We stress here that the converse is not true. The single centered black hole solution exists only when the discriminant $4mn - \ell^2$ is positive, and its degeneracy grows exponentially for large values of the discriminant. On the other hand, as we saw in §4, any Jacobi form with exponential growth necessarily has negative discriminant states. This means that the partition function for single centered black holes alone with fixed magnetic charge and varying electric charge cannot have good modular properties. Indeed the partition function ψ^{F}_m does contain negative discriminant states for example for $n < m$. These low-lying states do not correspond to single centered black holes [31, 108].

We now turn to the physical interpretation of the polar part $\psi^{\mathrm{P}}_m(\tau, z)$ of the meromorphic Jacobi form defined in (8.44). As mentioned in the introduction, the only multi-centered gravitational configurations that contribute to the supersymmetric index of quarter-BPS dyons in our $\mathcal{N} = 4$ string theory have exactly two centers [38] each of which is a half-BPS state. At a wall, one of these two-centered configurations decays. A basic decay at a wall-crossing is when one center is purely electric and the other is purely magnetic. Other decays are related to this basic decay by duality [105]. For this two-centered configuration, with one electric and one magnetic center, the indexed partition function after summing over the electric charges is:

$$p_{24}(m+1) \cdot \frac{1}{\eta^{24}(\tau)} \cdot \sum_{\ell>0} \ell\, y^{\ell}\,, \tag{11.18}$$

where the first factor is the degeneracy of the half-BPS magnetic center, the second factor is the partition function of the electric center [39], and the third factor is the partition function counting the dimensions of the $SU(2)$ multiplets with angular momentum $(\ell-1)/2$ contained in the electromagnetic field generated by the two centers [46] for which $N \cdot M = \ell$. The third factor is an expansion of the function

$$\frac{y}{(1-y)^2} \tag{11.19}$$

in the range $|q| < |y| < 1$ (see equation (9.7)). Here we already see the basic wall-crossing of the theory because different ways of expanding this meromorphic function for $|y| < 1$ and $|y| > 1$ will give different degeneracies. Other wall-crossings are related to this basic wall-crossing by S-duality which as we have explained is nothing but spectral flow for a given fixed m. Thus the full partition function that captures all wall-crossings is obtained by 'averaging' over the spectral-flow images of (11.19). This averaging is nothing but the averaging operation defined in (9.3) and (9.4)

$$\mathrm{Av}^{(m)}\Big[\frac{y}{(y-1)^2}\Big] \;=\; \frac{q^{ms^2+s}y^{2ms+1}}{(1-q^s y)^2} \tag{11.20}$$

It is now easy to see that this averaging over spectral-flow images gives exactly the polar part ψ_m^{P} of the meromorphic Jacobi form ψ_m:

$$\frac{p_{24}(m+1)}{\eta^{24}(\tau)} \sum_{s\in\mathbb{Z}} \frac{q^{ms^2+s}y^{2ms+1}}{(1-q^s y)^2} \;=\; \psi_m^{\mathrm{P}}(\tau,z)\,. \tag{11.21}$$

This follows from the fact that the poles of $\psi_m(\tau,z)$ are exactly at $z \in \mathbb{Z}\tau + \mathbb{Z}$ and the residue at $z = 0$ is

$$\frac{p_{24}(m+1)}{\eta^{24}(\tau)}\,.$$

as mentioned in (1.5).

We thus see that the decomposition theorem (8.3) as applied to our dyon partition function $\psi_m(\tau,z)$ has a natural physical interpretation. It is simply the statement that the full partition function is a sum of its immortal and decaying constituents:

$$\psi_m(\tau,z) \;=\; \psi_m^{\mathrm{F}}(\tau,z) \,+\, \psi_m^{\mathrm{P}}(\tau,z)\,. \tag{11.22}$$

The nontrivial part is of course the implication of such a decomposition for modularity. By separating part of the function ψ_m, we have evidently broken modularity, and $\psi_m^{\mathrm{F}}(\tau,z)$ is not a

Jacobi form any more. However, the decomposition theorem (8.3) guarantees that ψ^{F}_m still has a very special modular behavior in that it is a mock Jacobi form. The mock behavior can be summarized by the following partial differential equation obeyed by its completion:

$$\tau_2^{3/2}\,\frac{\partial}{\partial\overline{\tau}}\,\widehat{\psi}^{\mathrm{F}}_m(\tau,z)\;=\;\sqrt{\frac{m}{8\pi i}}\;\frac{p_{24}(m+1)}{\Delta(\tau)}\sum_{\ell\,\mathrm{mod}\,(2m)}\overline{\vartheta_{m,\ell}(\tau,0)}\,\vartheta_{m,\ell}(\tau,z)\,. \tag{11.23}$$

Note that ψ^{P}_m also admits a completion $\widehat{\psi^{\mathrm{P}}_m}$ which is modular. However, the function $\widehat{\psi^{\mathrm{P}}_m}$, unlike $\widehat{\psi^{\mathrm{F}}_m}$, is meromorphic, and therefore has wall-crossings. We refer the reader to the summary at the end of §8.3 for a detailed discussion of the relationship between meromorphy (in z), holomorphy (in τ) and modularity.

Let us now discuss the implications of these results for AdS_2/CFT_1 holography. The near horizon geometry of a supersymmetric black hole is a two-dimensional anti de Sitter space AdS_2 and the dual theory is a one-dimensional conformal field theory CFT_1. The partition function of this CFT_1 is nothing but the integer $d^*(n,\ell,m)$ which gives the number of microstates of the black hole [106, 107]. Our results show that that these integers are the Fourier coefficients of a mock modular form and as a result there is a hidden modular symmetry. Such a modular symmetry would make it possible to use powerful techniques such as the Hardy-Ramanujan-Rademacher expansion to process the integers $d^*(n,\ell,m)$ into a form that can be more readily identified with a bulk partition function of string theory in AdS_2. There has been recent progress in evaluating the bulk partition function using localization techniques which makes it possible to study the subleading nonperturbative corrections in the Rademacher expansion in a systematic way [34, 35, 70, 96]. It would be interesting to see if the additional terms in the Rademacher expansion arising from the 'mock' modular nature can be given a physical bulk interpretation.

11.2 Meromorphy, mock modularity, and noncompactness

From the mathematical point of view, we have seen how the mock modularity of ψ^F_m is a consequence of meromorphy of the Jacobi form ψ_m. In the $\mathcal{N}=4$ example above, we can understand the physical origin of this meromorphy as coming from the non-compactness of the underlying brane dynamics. To see this more clearly, we write the meromorphic Jacobi form $\psi_m(\tau,z)$ as

$$\psi_m(\tau,z)=\frac{1}{\varphi_{-2,1}(\tau,z)}\,\frac{1}{\Delta(\tau)}\;\chi(\tau,z;\mathrm{Sym}^{m+1}(K3))\,. \tag{11.24}$$

The three factors in the above equations have the following physical interpretation [42]. The function $\psi_m(\tau,z)$ can be viewed as an indexed partition function of the following $(0,4)$ SCFT

$$\sigma(\mathrm{TN})\times\sigma_L(\widetilde{K}\text{-}n)\times\sigma(\mathrm{Sym}^{m+1}(K3))\,. \tag{11.25}$$

Each CFT factor has a natural interpretation from the microscopic derivation of the dyon counting formula [42] using the 4d-5d lift [60]. The $\sigma_L(\widetilde{K}\text{-}n)$ factor is the purely left-moving bosonic CFT associated with $\widetilde{K}$-n bound states of the Kaluza-Klein monopole with momentum. The $\sigma(\mathrm{TN})$ factor is a $(4,4)$ SCFT associated with the motion of the center of mass of motion in the Kaluza-Klein background in going from five dimensions to four dimensions in the 4d-5d lift. The $\sigma(\mathrm{Sym}^{m+1}(K3))$ is the $(4,4)$ SCFT that is associated with the five-dimensional Strominger-Vafa $Q1$-$Q5$-n system. Note that the SCFT $\sigma(\mathrm{TN})$ has a *noncompact* target space and the double pole of $1/\varphi_{-2,1}(\tau, z)$ can be traced back to this noncompactness.

Going over from the partition function to the Fourier coefficients corresponds to going over from the canonical ensemble with fixed z to the micro-canonical ensemble with fixed ℓ. The Fourier coefficient is supposed to count the number of right-moving ground-states in this fixed charge sector. Right-moving states with nonzero energy cancel in pairs in this index. Now, counting ground states for a compact theory is a well-defined operation. However, when the target-space is noncompact, the wavefunctions are not properly normalized without putting an infrared regulator in the target space. In the present context, the essential physics of the noncompactness is captured by the point particle motion of the center of mass in the Taub-NUT geometry. A simple regulator can be introduced by turning on the axion field which according to (2.22) corresponds in the Type-IIB frame to taking the modulus $U_B = U_{1B} + iU_{2B}$ to have a small nonzero real part U_{1B} [42]. This introduces a nonzero potential on the Taub-NUT geometry proportional to $U_{1B}^2 U_{2B}$ that regulates the infrared behavior to give a well defined counting problem [62, 100]. However, the index thus computed depends on the sign of U_{1B} and jumps as U_{1B} goes from being slightly negative to slightly positive. Indeed the partition function for these modes is precisely the function

$$\frac{y}{(1-y)^2} \tag{11.26}$$

which can be expanded either in power of y or in powers of y^{-1} depending on whether U_{1B} is positive or negative. For ℓ positive, there are ℓ normalizable states when $U_{1B} < 0$ and none when $U_{1B} > 0$. This is the essential source of the wall-crossing behavior in this context. We should emphasize that even though turning on U_{1B} appears to make the problem effectively compact, this is true only up to some maximal ℓ_{max} which scales as $U_{1B}^2 U_{2B}$ [62, 100]. For a given ℓ positive, one can always make the problem well-defined with a compact target space by choosing the moduli so that ℓ_{max} is larger than the ℓ of interest. The partition function in that case is well-defined without any poles and goes as

$$y + 2y^2 + \ldots + \ldots + \ell_{max} y^{\ell_{max}}\,. \tag{11.27}$$

However, if one wishes to compute the canonical partition function that sums over all ℓ, one has to essentially take $U_{1B}^2 U_{2B}$ to go to infinity which is what leads to the pole at $y = 1$ in

(11.26). This in essence is the origin of the meromorphy of the canonical counting function of the asymptotic states, and therefore indirectly also of the mock modularity of the immortal near-horizon counting function.

As we have seen, this meromorphy of the asymptotic counting function ψ_m implies that the counting function of the near-horizon immortal states cannot be both modular and holomorphic at the same time. One can either have the holomorphic function ψ_m^{F} which is not quite modular, or a modular function $\widehat{\psi}_m^{\mathrm{F}}$ which is not quite holomorphic. From this point of view, the shadow can be regarded as a 'modular anomaly' in ψ_m^{F} or as a 'holomorphic anomaly' in the completion $\widehat{\psi}_m^{\mathrm{F}}$.

A natural question that arises from this discussion is if the mock Jacobi form and the holomorphic anomaly can be given a direct physical interpretation rather than arrive at it via a meromorphic Jacobi form as we have done. To address this question, it is useful to view the problem from the perspective of AdS_3 holography in M-theory as we now discuss.

11.3 The M-Theory limit and AdS_3 holography

From the perspective of AdS_2/CFT_1 holography discussed above, there is no *a priori* reason why the black hole degeneracy should have anything to do with modularity. It is only when we can view the black hole as an excitation of a black string that we can have an *a priori* physical expectation of modularity. This is because the near horizon geometry of a black string is a three-dimensional anti de Sitter space AdS_3. The Euclidean thermal AdS_3 has a conformal torus as a boundary whose complex structure parameter is τ. The partition function of the dual boundary conformal field theory CFT_2 depends on on τ. In this framework the modular symmetry is identified with the $SL(2,\mathbb{Z})$ mapping class group of the boundary torus [86], whereas the elliptic symmetry is identified with large gauge transformations of the 3-form field [44] as we have already discussed in §11.1.

We now describe how the quarter-BPS black holes considered thus far can be regarded as excitations of a black string in an appropriate M-theory frame. This will enable us to make contact with AdS_3/CFT_2 holography. The M-theory framework is natural also from the point of view of generalizing these considerations to BPS black holes in $\mathcal{N}=2$ compactifications on a general Calabi-Yau manifold X_6.

To go to the M-theory frame, we consider the T^2 to be a product of two circles $S^1 \times \tilde{S}^1$ and consider the charge configuration (2.15) in the Type-IIB frame. We first T-dualize the $\tilde{S}$ circle to the $\hat{S}^1$ circle to go to the Type-IIA frame and then use mirror symmetry of $K3$ to map the Type-IIB configuration (2.15) to a Type-IIA configuration of D4-F1-n-NS5 charges. We then lift it to M-theory to obtain a charge configuration consisting of M5-M2 bound states carrying momentum along S^1. This configuration now has M5-branes wrapping $D \times S^1$ where D is a divisor in $K3 \times \hat{T}^2$, an M2-brane wrapping the $\hat{T}^2 = \hat{S}^1 \times S^1_m$, with some momentum along S^1. This is a precisely a configuration of the kind considered by Maldacena, Strominger, and Witten [87]. Since all M5-branes wrap S^1 and the momentum is along S^1, it is natural flip S^1_m and S^1 and regard the S^1 as the new M-circle. In the limit when the radius R of the new M-circle S^1 is very large, the low-lying excitations of this M5-brane are described by the MSW string wrapping the M-circle. To write the final configuration in the M-theory frame, let C^1 to be the homology 2-cycle of the $\hat{T}^2$, and let C^2 and C^3 be two 2-cycles in $K3$ with intersection matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{11.28}$$

The intersection number of these three 2-cycles is

$$\int_{K3\times\hat{T}^2} C^1 \wedge C^2 \wedge C^3 = 1\,. \tag{11.29}$$

Let $\{D_a\}$ be the 4-cycles dual to $\{C^a\}$, $D_a \cap C^b = \delta^b_a$. In this basis, the M5-brane is wrapping the 4-cycle

$$D = \sum_{a=1}^{3} p^a D_a$$

with $p^1 = \tilde{K}$, $p^2 = Q_1$ and $p^3 = Q_5$. The M2-brane charges are $q_2 = q_3 = 0$ and $q_1 = \tilde{n}$, and M-momentum or equivalently the D0-brane charge from the Type-IIA perspective is $q_0 = n$. Using this notation we see that

- The electric charges in (2.15) map to n units of momentum along the M-circle S^1, and $\tilde{K}$ M5-branes wrapping $D_1 \times S^1$.
- The magnetic charges in (2.15) map to Q_1 M5-branes wrapping $D_2 \times S^1$, Q_5 M5-branes wrapping $D_3 \times S^1$, and $\tilde{n}$ M2-branes wrapping $\hat{T}^2$.

In summary, the charge configuration of our interest is

$$\Gamma = \begin{bmatrix} N \\ M \end{bmatrix} = \begin{bmatrix} 0, & n; & 0, & \tilde{K} \\ Q_1, & \tilde{n}; & Q_5, & 0 \end{bmatrix}_B = \begin{bmatrix} 0, & q_0; & 0, & p^1 \\ p^2, & q_1; & p^3, & 0 \end{bmatrix}_M \tag{11.30}$$

Reduction of this charge configurations along S^1 gives a configuration in Type-IIA frame consisting of only D4-D2-D0-branes.

Since the effective MSW string wraps the M-circle, we have to take the radius R of this circle to be large keeping other scales fixed to obtain an AdS_3 near-horizon geometry of a long black string. This implies that in the original heterotic frame in which we have labeled the charges (2.3), the radius R goes to infinity keeping other scales and moduli fixed. Since the 2-torus in the heterotic frame is also a simple product of two circles $S^1 \times \tilde{S}^1$, the moduli in (2.20) take the form

$$G_{ij} = \begin{pmatrix} \tilde{R}^2 & 0 \\ 0 & R^2 \end{pmatrix} , \quad B_{ij} = 0 \tag{11.31}$$

which implies that for the three moduli S, T, U we have

$$S \; = \; a + i\frac{R\tilde{R}}{g_6^2}\,, \qquad T \; = \; iR\tilde{R}\,, \qquad U = i\,\frac{R}{\tilde{R}}\,, \tag{11.32}$$

where g_6 is the six-dimensional string coupling constant of heterotic compactified on T^4 and we have allowed a nonzero axion field a. This means in particular that

$$S_2 \sim R, \quad T_2 \sim R, \quad U_2 \sim R \tag{11.33}$$

and for the charge configuration (11.30), the central charge matrix scales as

$$\mathcal{Z} \sim \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{R}\begin{pmatrix} 2n & \ell + ag_6^2/\tilde{R} \\ \ell + ag_6^2/\tilde{R} & 2m + (Q_1^2/\tilde{R}^2 + Q_5^2\tilde{R}^2) \end{pmatrix} . \tag{11.34}$$

From the contour prescription (6.10), we conclude that the imaginary parts for the Fourier integral scale as

$$\mathrm{Im}(\sigma) \sim R \tag{11.35}$$

$$\mathrm{Im}(\tau) \sim 1/R \tag{11.36}$$

$$\mathrm{Im}(z) \sim 1/R\,. \tag{11.37}$$

Therefore, in the region of the moduli space corresponding to the M-theory limit, $p := \exp(2\pi i\sigma)$ is becoming much smaller compared to $q := \exp(2\pi i\tau)$ and $y := \exp(2\pi i z)$. Hence one must first expand around $p = 0$:

$$\frac{1}{\Phi_{10}(\Omega)} \; = \; \sum_{m=-1}^{\infty} p^m \psi_m(\tau, z) \; . \tag{11.38}$$

Since the function $\Phi_{10}(\Omega)$ has double zeros at $z = 0$, the function $\psi_m(\tau, z)$ is meromorphic with double poles at $z = 0$ and its translates.

We conclude that the function $\psi_m(\tau, z)$ can be interpreted as the asymptotic partition function for counting the BPS excitation of this MSW M5-brane for fixed value of the magnetic charge $M^2/2 = m$ and is a *meromorphic* Jacobi form of weight -10 and index m. This is the essential new ingredient compared to other situations encountered earlier such as the partition function of the D1-D5 string [118] which is a *holomorphic* weak Jacobi form[28].

Note that $\psi_m(\tau, z)$ is the asymptotic counting function in the canonical ensemble for a fixed chemical potential z where one sums over all M2-brane charges ℓ for a fixed chemical potential z. To obtain the micro-canonical partition function for a fixed M2-brane charge, one has to perform the Fourier integral in z. There is an infinity of double poles in the z plane at $z = \alpha\tau$ for $\alpha \in \mathbb{Z}$. The z contour can be made to lie between the poles at $(\alpha - 1)\tau$ and $\alpha\tau$ depending on how $\mathrm{Im}(z)$ compares with $\mathrm{Im}(\tau)$ given the $O(1/R)$ terms in the central charge matrix (11.34). It is easy to see that by varying the moduli, the contour can be made to cross all poles. Thus, there are an infinity of walls still accessible in the regime when the radius of the M-circle becomes large.

---

[28] Here we mean holomorphic or meromorphic in the z variable that is conjugate to ℓ. This should not be confused with the nomenclature used in §4 while discussing growth conditions (4.5).

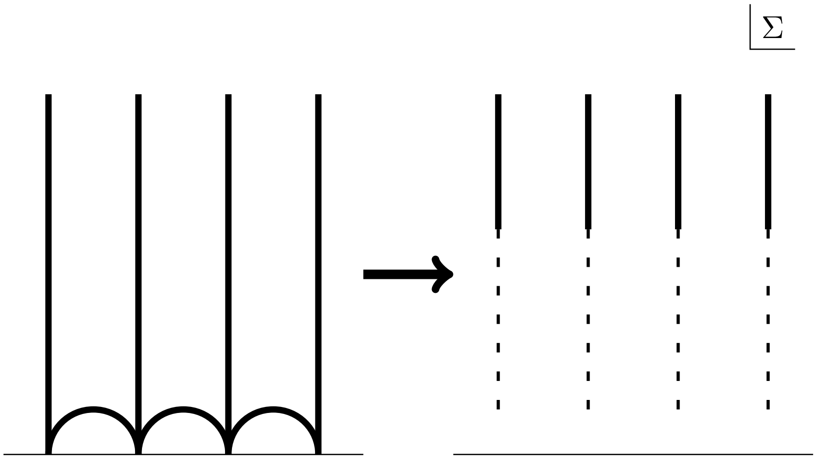


Figure 1: The diagram on the left shows the moduli space projected onto the Σ upper-half plane divided into chambers separated by walls where a quarter-BPS state decays into two half-BPS states. The M-theory limit, shown on the right, corresponds to taking $\mathrm{Im}(\Sigma)$ to be very large. In this limit several walls are no longer visible. There are still an infinite number of walls which can be crossed by varying $\mathrm{Re}(\Sigma)$. Spectral flow transformation maps one chamber to another chamber.

To display this graphically it is useful to define a complex scalar field $\Sigma = \Sigma_1 + i\Sigma_2$ by

$$\mathcal{Z} \sim \frac{1}{\Sigma_2}\begin{pmatrix} |\Sigma|^2 & \Sigma_1 \\ \Sigma_1 & 1 \end{pmatrix} . \tag{11.39}$$

For fixed charges, the MSW region which corresponds to taking Σ_2 large with Σ_1 fixed corresponding to vanishing axion. Varying Σ_1 allows one to access the infinite number of chambers separated by walls as shown in Fig. 1. Degeneracies in these different chambers can be obtained simply by choosing the Fourier contour in z for an appropriate value of α. Note that the field Σ_1 depends both on the moduli and on the charges and thus varies as moduli are varied for fixed charges or as the charges are varied for fixed moduli. In particular, even for fixed values of moduli, one can cross all walls by varying ℓ appropriately.

With this embedding into M-theory, we now outline what we regard as a consistent physical picture to interpret $\widehat{\psi}^F_m$ even though we do not yet fully understand the details. In the limit of large radius R, we effectively have a five-dimensional theory obtained from compactifying M-theory on $K3 \times \hat{T}^2$. For a given a set of M5-brane charges $\{p^i\}$ at asymptotic infinity, the five-dimensional supergravity admits a single-centered black string solution carrying magnetic charges $\{p^i\}$. The near horizon geometry of this solution is an AdS_3 geometry where all vector multiplet moduli of the five-dimensional supergravity are fixed to their attractor values. Following the usual reasoning of holography, we expect that the superconformal field theory describing the low energy excitations of the M5-brane will be dual to this near horizon AdS_3. The indexed partition function of this SCFT is expected to describe the BPS gravitational configurations inside the AdS_3.

What is the indexed partition function of this SCFT? It cannot possibly be the meromorphic Jacobi form ψ_m that exhibits many wall-crossings, because all vector-multiplet moduli are fixed in the attractor AdS_3 geometry and no wall-crossings are possible. It cannot possibly be ψ^F_m, because it is not modular as would be expected from the symmetry under the $SL(2,\mathbb{Z})$ mapping class group of the boundary torus. We propose that the most natural object to identify with the indexed partition function of the $SCFT_2$ dual to the AdS_3 is the modular completion $\widehat{\psi}^F_m(\tau, z)$. It satisfies (in a rather nontrivial way) both requirements of being properly modular and not having any wall-crossings. The non-holomorphy is not something that we would naively expect for the indexed partition function but this can be a reflection of noncompactness as we comment upon later. The quarter-BPS black holes discussed in this paper are excitations inside this single AdS_3 represented as BTZ black holes and their degeneracies should then be captured by the Fourier coefficients of $\psi^F_m(\tau, z)$ which is indeed the case for $n > m$.

This *microscopic* picture is consistent with the *macroscopic* analysis of multi-centered black string solutions [45] of five dimensional supergravity. These solutions can be obtained by lifting multi-centered black hole solutions of four-dimensional supergravity. According to this supergravity analysis, given a set of asymptotic magnetic charges $\{p^i\}$ of a wrapped M5-brane, the single-centered AdS_3 throat described above is not the only allowed solution [45]. There exist multi-centered solutions with multiple AdS_3 throats with the same total asymptotic charge. Now, as we have seen from the microscopic analysis, there are still infinite chambers separated by walls visible in the M-theory limit as shown in Fig 1. These chambers can thus correspond to chambers in which different multi-string configurations exist as stable solutions. The wall-crossings can result not from decays of multi-centered black holes inside a single AdS_3 throat but from the decays of these multi-string solutions. Thus, it is reasonable to identify $\widehat{\psi}^F_m$, which has no wall-crossings, with the (indexed) partition function of the CFT dual to the single AdS_3 throat.

11.4 Open problems and comments

The interpretation of $\widehat{\psi}^F_m(\tau, z)$ proposed above raises a number of interesting questions. The essentially new feature of our proposal is the fact that the modular partition function $\widehat{\psi}^F_m(\tau, z)$ is nonholomorphic and correspondingly the holomorphic counting function ψ_m^F is mock modular. It would clearly be important to understand the physical origin of this nonholomorphy and mock modularity from the boundary and bulk perspectives.

- *Mock modularity from the boundary perspective*:

  We have seen in §11.2 that the mock modularity of ψ_m^F is tied to the noncompactness of the asymptotic SCFT. A more interesting question is to understand this mock modularity directly from the point of view of the near horizon SCFT. It is natural to assume that the near horizon SCFT is also noncompact and this noncompactness is what is responsible for the mock modularity. The connection between mock modularity/nonholomorphy and noncompactness has been noted earlier in physics. For example, the partition function of topological $SO(3)$ $\mathcal{N} = 4$ Yang-Mills theory on $\mathbb{CP}^2$ is a mock modular form and it was suggested in [120] that it is related to the noncompactness of flat directions in field space. The nonholomorphy of the elliptic genus has also been noted in the context of superconformal field theories [54, 119] with the noncompact $SL(2,\mathbb{R})/U(1)$ target space.

  In this context, the nonholomorphy can be understood in complete detail [119, 4] as a consequence of the continuum of spectrum of the noncompact SCFT. It would be interesting to see if the holomorphic anomaly (11.23) of the completion of $\widehat{\psi}_m^F$ can be understood from such a path integral perspective of the conformal field theory.

  In the framework of AdS_2 holography, the quantity of interest is the integer $d^*(n, m, \ell)$ which is identified with the partition function of the boundary CFT. Using the fact that the completion $\widehat{\psi}_m^F(\tau, z)$ is modular, it should be possible to develop a Rademacher-like expansion for the $d^*(n, m, \ell)$ along the lines of [13, 14, 15]. This method has been used to prove that the black hole degeneracies are positive integers [19].

- *Mock modularity from the bulk perspective*:

  If mock modularity is indeed a consequence of the noncompactness of the boundary CFT, what are the implications of this noncompactness in the bulk AdS_3? In particular, if the spectrum of conformal dimensions is continuous, then the bulk theory must have a continuum of masses. Perhaps these are to be identified with some extended long-string like excitations. It would be interesting to see if the mock modularity of the elliptic genus can be understood in terms of such bulk excitations. Indeed, recent investigations seem to indicate that the bulk string partition function in string theories [71] based on the non-compact $SL(2,\mathbb{R})/U(1)$ CFT [64, 94] show a mock modular behavior.

  The M-theory limit corresponds to the attractor contour for $n > m$ and $0 \leq \ell < 2m$ but the converse is not necessarily true, as explained in §11.1. For $n \leq m$, we have the possibility that there will be multi-centered configurations. In particular, if $n \leq m$, then even if $\ell < 2m$, it is possible to have states with non-positive discriminant ($\Delta = 4mn - \ell^2 \leq 0$). Since, the discriminant must be positive for a black hole horizon, such states cannot possibly correspond to single-centered black holes and must be realized in supergravity as multi-centered configurations if there are smooth solutions corresponding to them at all. Hence, if $\widehat{\psi}^F_m$ is indeed the elliptic genus of a single-centered AdS_3, then there must be supergravity configurations corresponding to these states for $n \leq m$ that fit inside a single AdS_3.

This microscopic prediction is plausible from the supergravity perspective. According to [45], all two-centered black hole solutions where each center is a D4-D2 bound state, are scaled out of a single AdS_3 throat. However, the two-centered solutions that have one center carrying D6-brane charge and the other carrying anti-D6-brane charge can continue to exist as stable solutions inside a single AdS_3. Such configurations can give the supergravity realization of the negative discriminant states. Further analysis is required to verify these consequences of our proposed interpretation of $\widehat{\psi}^F_m$.

- *Generalization to $\mathcal{N} = 2$ compactifications*:

One motivation for the present investigation is to abstract some general lessons which could be applicable to the more challenging case of dyons in $\mathcal{N} = 2$ compactifications on a general Calabi-Yau three-fold. In the $\mathcal{N} = 4$ case, the mock modularity of ψ^F_m was a consequence of meromorphy of ψ_m. As we have seen in §11.2, this meromorphy is in turn a consequence of noncompactness of Taub-NUT factor $\sigma(\mathrm{TN})$ in the asymptotic SCFT (11.25)[29]. This suggests that the SCFT dual to the near horizon AdS_3 is also noncompact. In the $\mathcal{N} = 2$ context also, there is no a priori reason to expect that the MSW SCFT will be compact in general everywhere in the moduli space. If this SCFT is noncompact, then our results suggest that mock Jacobi forms would provide the right framework for making the holographic $SL(2,\mathbb{Z})$ symmetry manifest for the indexed partition function considered, for example, in [49, 44, 89]. In the $\mathcal{N} = 4$ case, the mock modularity is closely correlated with the wall-crossing phenomenon. The wall-crossing phenomenon in $\mathcal{N} = 2$ case corresponds to jumps in the Donaldson-Thomas invariants and is known to be much more complicated [48, 78, 76, 88]. It would be interesting though to define the analog of the counting function for the immortal degeneracies in this case. Since these degeneracies do not change under wall-crossing, such a counting function is likely to have an interesting mathematical interpretation. Moreover, it is expected to have nice (mock) modular properties from the considerations of holography discussed above.

---

[29] Note that this is not to be confused with $\mathbb{R}^3$ factor corresponding to the center of mass motion. It corresponds to the relative motion of the KK-monopole and the D1D5 system [42] and thus would belong to what is called the "entropic" factor [92] for the MSW string in the M-theory frame.

- *Mock modularity and indefinite theta series*:

  Mock modular forms appear in Zwegers's work as members of three quite different families of functions:

  1. Fourier coefficients of meromorphic Jacobi forms,
  2. Appell-Lerch sums,
  3. Indefinite theta series.

  We have seen that the first two families appear naturally in the context of black hole physics with natural physical interpretations. This suggests that the indefinite theta series could also play a role in this physical context. It would be interesting to develop this connection further. For some earlier work in this direction, see [90, 1].

Acknowledgments

We would like to acknowledge useful conversations with Kathrin Bringmann, Miranda Cheng, Jan de Boer, Rajesh Gopakumar, Jeff Harvey, Jan Manschot, Shiraz Minwalla, Boris Pioline, Jan Troost, Erik Verlinde, Sander Zwegers, and especially Ashoke Sen. We would also like to thank Alexander Weisse for technical help. It is a pleasure to thank the organizers of the 2007 workshop on 'Black Holes and Modular Forms' in Paris and of the 2008 workshop on 'Number Theory and Physics at Crossroads' in Banff for hospitality where this research was initiated. The work of A. D. was supported in part by the Excellence Chair of the Agence Nationale de la Recherche (ANR). The work of S. M. was supported by the ERC Advanced Grant no. 246974, *"Supersymmetry: a window to non-perturbative physics"*.
v2: The work of A. D. was conducted within the framework of the ILP LABEX (ANR-10-LABX-63) supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02. S. M. would like to thank the SFTC for support from Consolidated grant number ST/J002798/1.

A. Appendix: Tables of special mock Jacobi forms

This appendix contains three collections of tables. The first gives the first Fourier coefficients of the weight 2 mock Jacobi forms $\mathcal{Q}_M$ for all products of two primes less than 50, the second gives coefficients of the primitive mock Jacobi forms $\Phi^0_{1,m}$ for all known cases having optimal growth, and the third gives coefficients of the weight 1 mock Jacobi forms $\mathcal{Q}_M$ for all products of three primes less than 150.

A.1 Table of $\mathcal{Q}_M$ (weight 2 case)

For each M with $\mu(M) = 1$ with $1 < M < 50$ we tabulate the value of c_M and the Fourier coefficients $c(\mathcal{F}_M; n, r_{\min})$ for n ranging from 0 to some limit and all values of $r_{\min}$, (defined as the minimal positive representatives prime to M of their squares modulo $4M$). Here c_M is the smallest positive rational number such that $\mathcal{F}_M = -c_M\mathcal{Q}_M$ has integral coefficients, as defined in (9.52). The asterisk before a value of $r_{\min}$ means that the corresponding row contains at least one negative non-polar coefficient.

M = 6 $c_M = 12$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12
1	−1	35	130	273	595	1001	1885	2925	4886	7410	11466	16660	24955

M = 10 $c_M = 6$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	−1	21	63	112	207	306	511	693	1071	1442	2037	2709	3766	4788
3	0	9	35	57	126	154	315	378	625	819	1233	1491	2268	2772

M = 14 $c_M = 4$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	−1	15	42	65	120	150	255	312	465	575	819	975	1355	1605
3	0	10	30	51	85	120	195	230	360	465	598	765	1065	1235
5	0	3	15	20	51	45	113	105	195	215	345	348	595	615

M = 15 $c_M = 3$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−1	14	29	54	83	128	172	230	331	430	537	726	924	1116	1409
2	0	4	24	28	58	52	136	120	224	232	376	368	626	632	920
4	0	10	28	40	88	98	176	208	304	386	552	632	888	1050	1372
7	0	2	9	26	36	50	84	116	155	226	269	340	488	580	712

M = 21 $c_M = 2$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	8	15	32	40	56	79	96	135	176	191	248	336	368	434	576
2	0	8	26	28	56	64	112	104	180	184	272	300	416	432	606	636
4	0	2	12	8	32	14	60	40	80	64	148	96	208	166	272	268
5	0	8	17	32	49	64	85	120	150	200	241	288	374	448	534	656
8	0	2	12	20	32	38	76	72	112	138	176	200	304	302	408	472
11	0	0	3	8	11	24	26	32	59	64	74	112	138	144	189	248

M = 22 $c_M = 12$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	−5	49	132	176	308	357	572	604	951	1061	1391	1572	2227	2325
3	0	44	103	169	272	301	536	580	771	976	1307	1380	2004	2121
5	0	25	91	116	223	229	427	405	716	700	1071	1121	1608	1633
7	0	13	48	79	127	180	272	272	452	549	632	764	1120	1133
9	0	1	23	24	91	25	180	92	247	205	403	229	715	453

M = 26 $c_M = 2$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	9	21	30	51	54	87	96	138	149	207	219	308	315	399	468
3	0	7	21	27	43	51	90	76	132	144	183	201	297	283	393	423
5	0	6	15	23	39	42	69	75	102	128	165	165	243	258	324	371
7	0	3	12	14	33	27	57	51	90	84	132	126	201	183	261	267
9	0	1	6	12	15	21	36	34	57	63	70	93	138	118	162	210
11	0	0	3	2	12	3	26	6	30	24	57	15	86	45	87	81

M = 33 $c_M = 6$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−5	33	55	127	127	160	242	259	347	468	468	534	749	826	903
2	0	21	66	43	142	120	252	174	340	262	526	394	658	580	844
4	0	24	88	88	165	154	286	244	418	376	572	596	869	728	1122
5	0	14	36	47	80	113	113	160	215	259	229	372	460	386	507
7	0	33	44	99	143	167	209	308	365	440	506	541	748	847	915
* 8	0	−4	29	18	62	−4	146	−8	124	76	234	54	336	116	380
10	0	11	50	66	116	110	248	220	292	358	474	435	738	622	892
13	0	0	21	54	77	120	131	153	230	273	317	405	482	548	613
16	0	0	3	22	58	36	94	124	127	146	248	182	350	380	408
19	0	0	0	7	18	7	40	40	58	113	58	113	186	153	120

M = 34 $c_M = 3$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−2	14	36	43	69	82	123	117	183	195	253	252	364	352	481
3	0	13	29	44	77	61	125	116	164	185	244	234	381	334	427
5	0	11	28	39	59	67	106	107	162	157	214	242	321	306	391
7	0	7	26	28	53	54	100	74	149	144	194	177	292	252	378
9	0	5	14	27	41	43	79	72	92	125	165	141	236	234	262
11	0	1	13	12	36	26	57	43	100	69	124	100	196	140	236
13	0	0	5	12	16	18	34	32	54	65	53	85	128	80	139
* 15	0	0	2	1	13	−2	35	0	29	14	59	−2	94	28	72

M = 35 $c_M = 1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	10	14	18	31	44	49	60	76	86	98	134	147	168	200	226
* 2	0	−2	8	4	14	−2	24	8	30	12	56	0	70	22	72	44
3	0	2	6	16	11	18	24	26	43	50	44	58	84	70	93	124
4	0	6	10	16	28	26	46	40	56	70	90	96	128	128	152	176
6	0	6	16	16	30	32	56	52	78	80	104	108	152	150	208	216
8	0	2	10	8	20	12	42	24	52	40	74	48	104	90	122	120
9	0	2	3	8	14	18	17	16	30	42	39	58	69	60	71	100
11	0	2	5	16	22	20	30	42	52	68	67	84	114	118	121	166
13	0	0	5	10	9	18	24	34	33	42	59	52	79	92	99	126
16	0	0	2	0	12	12	16	8	28	18	50	40	56	36	88	56
18	0	0	0	4	6	6	24	8	30	24	32	40	64	46	80	80
23	0	0	0	0	1	2	5	8	7	16	11	8	33	26	19	50

M = 38 $c_M = 4$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−3	21	44	59	103	97	159	149	238	240	322	300	465	445	540
3	0	16	46	62	80	102	166	138	226	231	287	318	449	391	559
5	0	15	43	48	91	75	151	148	194	208	303	269	427	354	512
7	0	14	29	43	82	83	119	117	179	197	256	252	360	358	433
9	0	6	30	36	61	55	124	77	179	165	201	193	360	256	404
11	0	4	16	31	47	45	88	91	108	136	168	155	257	254	300
13	0	0	14	14	40	26	72	45	103	64	169	105	207	137	249
15	0	0	5	11	14	29	43	25	57	75	58	76	160	88	133
*17	0	0	1	1	20	−7	33	−5	47	21	60	−14	109	15	83

M = 39 $c_M = 1$

n / $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	6	8	16	22	26	31	42	47	60	72	72	94	110	115	154
2	0	4	14	16	26	20	52	32	64	60	86	76	128	100	148	152
4	0	2	12	6	20	16	36	22	48	32	68	46	96	66	124	86
5	0	6	11	18	22	34	37	44	64	66	71	102	116	120	137	170
7	0	2	3	10	12	12	16	18	28	42	28	40	64	44	57	88
8	0	4	12	12	28	22	44	40	56	62	88	74	120	108	156	148
* 10	0	0	4	0	10	−4	22	4	18	0	36	0	46	16	42	24
11	0	2	7	16	17	24	34	36	47	66	62	76	107	104	117	154
14	0	0	6	8	14	20	28	24	46	40	62	60	86	72	118	116
17	0	0	1	6	10	12	15	24	31	34	35	44	61	72	71	86
20	0	0	0	2	4	6	16	8	16	20	24	30	48	32	44	60
23	0	0	0	0	1	4	3	2	11	12	4	18	19	6	25	32

M = 46 $c_M = 12$

n / $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	−11	57	161	184	276	299	483	414	667	644	874	897	1219	1023
3	0	69	115	184	299	253	493	437	585	690	838	736	1288	1127
5	0	46	138	147	299	285	414	368	667	607	782	676	1205	1044
7	0	46	123	161	215	230	414	353	560	575	744	790	1043	905
9	0	30	92	138	230	168	421	329	467	444	766	559	1065	766
11	0	23	75	98	167	230	282	276	443	443	558	535	817	765
13	0	5	69	69	166	120	281	143	442	350	465	424	810	516
15	0	0	27	73	119	92	211	188	211	349	376	303	606	487
17	0	0	26	26	92	26	187	95	256	121	325	233	466	236
19	0	0	2	23	25	71	94	48	163	163	96	140	349	188
* 21	0	0	0	1	47	−22	93	−22	70	1	208	−68	324	−21

A.2 Table of $\Phi^0_{1,m}$ of optimal growth

We give the Fourier coefficients of the primitive weight 1 mock Jacobi form $\Phi^0_{1,m}$ for the fourteen known values of m for which it has optimal growth, in the same format as was used in the $k = 2$ case above. The number c is the factor relating the true value of $\Phi^0_{1,m}$ to the form whose coefficients have been tabulated, and which has been normalized to have integral Fourier coefficients with no common factor except for the values $C(\Delta = 0, r)$, which are allowed to have a denominator of 2. (This happens for the indices 9, 16, 18, and 25.)

$\mathbf{m = 2}$ $\quad c = 24$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11
1	-1	45	231	770	2277	5796	13915	30843	65550	132825	260568	494385

$\mathbf{m = 3}$ $\quad c = 12$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	-1	16	55	144	330	704	1397	2640	4819	8480	14509	24288	39765	63888
2	0	10	44	110	280	572	1200	2244	4180	7348	12772	21330	35288	56760

$\mathbf{m = 4}$ $\quad c = 8$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-1	7	21	43	94	168	308	525	882	1407	2255	3468	5306	7931	11766
2	0	8	24	56	112	216	392	672	1128	1840	2912	4536	6936	10416	15456
3	0	3	14	28	69	119	239	393	693	1106	1806	2772	4333	6468	9710

$\mathbf{m = 5}$ $\quad c = 6$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	4	9	20	35	60	104	164	255	396	590	864	1259	1800	2541	3560
2	0	5	15	26	54	90	156	244	396	590	905	1320	1934	2751	3924	5456
3	0	4	11	24	45	80	135	220	350	540	810	1204	1761	2524	3586	5040
4	0	1	6	10	25	36	76	110	189	280	446	636	970	1360	1980	2750

$\mathbf{m = 6}$ $\qquad c = 24$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-5	7	26	33	71	109	185	249	418	582	858	1184	1703	2291	3213
2	0	24	48	96	168	264	432	672	984	1464	2112	2976	4200	5808	7920
3	0	12	36	60	120	180	312	456	720	1020	1524	2124	3036	4140	5760
4	0	12	36	72	120	216	348	528	816	1212	1752	2520	3552	4920	6792
5	0	1	13	14	51	53	127	155	291	382	618	798	1256	1637	2369

$\mathbf{m = 7}$ $\qquad c = 4$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	2	3	5	10	15	21	34	48	65	94	129	175	237	312	413
2	0	2	6	10	16	24	40	54	84	116	164	222	310	406	552	722
3	0	3	6	11	18	29	45	66	95	137	192	264	361	486	650	862
4	0	2	6	8	18	24	42	58	90	122	180	240	338	448	612	794
5	0	1	3	7	11	18	28	41	63	91	125	177	245	328	441	590
6	0	0	2	2	6	6	16	18	32	40	66	80	126	156	224	286

$\mathbf{m = 8}$ $\qquad c = 4$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	1	2	4	6	7	14	18	25	35	48	63	87	110	146	190
2	0	2	4	6	10	16	22	32	46	62	86	116	152	202	264	340
3	0	2	5	7	13	18	29	38	58	77	108	141	195	250	333	424
4	0	2	4	8	12	18	28	40	56	80	108	144	196	258	336	440
5	0	1	4	5	11	14	24	32	50	63	94	122	170	215	294	371
6	0	0	2	4	6	10	16	22	32	46	62	86	116	152	202	264
7	0	0	1	1	4	3	9	8	17	20	33	36	60	70	101	124

$\mathbf{m = 9}$ $\qquad c = 3$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	0	2	2	3	6	8	10	15	20	27	36	46	58	78	98
2	0	2	3	4	8	9	15	20	27	37	51	63	87	109	142	178
3	0	2	3	6	9	12	18	26	36	48	63	84	111	144	183	234
4	0	1	4	5	9	13	21	25	39	50	70	90	122	151	202	251
5	0	2	3	6	8	12	18	26	34	48	65	84	112	146	186	240
6	0	$\frac{1}{2}$	3	3	7	9	15	18	30	36	54	66	93	115	156	192
7	0	0	1	2	4	6	9	12	19	26	33	46	62	78	103	134
8	0	0	1	1	3	1	6	5	10	10	18	19	33	35	52	60

m = 10 $c = 8$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−3	3	1	2	5	10	9	15	19	26	31	41	52	76	87	108
2	0	4	8	16	16	28	40	48	72	96	120	160	208	256	328	416
3	0	1	7	7	16	14	27	32	51	57	87	101	144	166	226	269
4	0	8	12	16	28	40	56	80	104	136	184	240	304	392	496	624
5	0	2	6	8	14	18	30	34	54	68	92	114	158	190	252	308
6	0	4	8	16	24	32	48	64	88	124	160	208	272	348	440	560
7	0	0	5	3	11	8	22	23	40	39	69	75	113	128	182	210
8	0	0	4	8	8	16	24	32	44	64	80	104	144	176	232	296
*9	0	0	1	−1	5	2	9	2	13	11	23	18	43	37	60	60

m = 12 $c = 8$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
*1	−3	−2	1	−1	−2	−2	1	−5	−1	−7	−5	−6	−5	−16	−10	−16
2	0	4	4	8	12	12	20	28	32	44	56	68	88	112	132	164
3	0	4	8	8	16	20	28	32	48	60	80	92	124	148	188	224
4	0	2	4	6	8	12	16	20	28	36	44	58	72	88	112	136
5	0	3	5	9	13	14	24	31	40	51	66	80	109	130	162	199
6	0	4	8	12	16	24	32	44	56	72	96	120	152	188	232	288
7	0	1	6	5	11	14	22	21	39	43	61	71	98	111	152	174
8	0	0	2	4	6	8	12	16	20	28	36	44	58	72	88	112
9	0	0	4	4	12	8	20	20	36	36	56	64	92	100	144	160
10	0	0	0	4	4	8	12	12	20	28	32	44	56	68	88	112
*11	0	0	0	−1	1	−3	1	−2	0	−5	0	−9	2	−8	−7	−14

m = 13 $c = 2$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	1	1	0	1	1	2	3	3	4	6	7	7	10	12	14
2	0	0	1	2	2	3	4	4	6	8	10	12	16	18	24	28
3	0	1	1	2	3	4	5	7	9	11	13	17	22	26	32	39
4	0	1	2	2	4	4	6	8	11	12	18	20	26	32	40	46
5	0	1	2	3	3	5	7	8	11	15	18	23	29	34	43	53
6	0	1	2	2	4	4	8	8	12	14	19	22	30	34	44	52
7	0	1	1	2	4	5	6	8	11	13	17	22	27	34	41	51
8	0	0	2	2	3	3	6	6	10	12	16	18	26	28	38	44
9	0	0	1	2	2	3	4	6	7	10	13	15	20	25	30	37
10	0	0	1	0	2	2	4	4	6	6	10	10	16	17	24	26
11	0	0	0	1	1	1	2	2	4	5	5	8	10	11	14	18
12	0	0	0	0	1	0	2	0	2	1	4	2	6	4	8	8

m = 16 $c = 2$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	0	0	0	1	0	1	2	1	1	2	3	3	4	5	5
2	0	0	1	1	1	1	2	2	3	4	4	5	7	7	9	11
3	0	1	1	1	2	2	3	3	4	5	7	7	9	11	13	15
4	0	1	1	1	2	3	3	4	5	6	8	9	11	14	16	19
5	0	0	1	2	2	2	4	4	6	7	9	10	14	14	19	22
6	0	1	1	2	2	3	4	5	6	8	9	11	14	17	20	24
7	0	1	2	1	3	3	4	5	7	7	10	11	15	17	21	23
8	0	$\frac{1}{2}$	1	2	2	3	4	4	6	8	9	12	14	16	20	24
9	0	0	1	1	2	2	4	4	6	6	9	10	14	14	19	22
10	0	0	1	1	2	2	3	4	5	6	8	9	11	14	17	20
11	0	0	1	1	2	2	3	2	5	5	7	7	11	11	15	16
12	0	0	0	1	1	1	2	3	3	4	5	6	8	9	11	14
13	0	0	0	0	1	0	2	1	3	2	4	4	6	6	9	9
14	0	0	0	0	0	1	1	1	1	2	2	3	4	4	5	7
15	0	0	0	0	1	0	1	0	1	1	2	0	3	1	3	2

m = 18 $c = 3$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
*1	−2	0	−1	0	−1	2	1	0	0	1	0	3	1	3	2	4
2	0	3	0	3	3	3	3	6	6	6	9	9	12	15	15	21
*3	0	−1	2	1	1	0	3	1	4	2	5	5	7	4	9	8
4	0	0	3	3	3	6	6	6	12	12	15	18	21	24	30	36
5	0	2	1	1	2	3	3	3	4	6	5	6	9	11	12	13
6	0	3	3	3	6	6	9	12	12	15	18	24	27	33	39	45
*7	0	−1	1	0	3	1	3	2	6	4	8	7	11	8	15	15
8	0	3	3	6	6	6	9	12	12	18	21	24	30	36	39	48
9	0	0	1	1	2	2	3	3	5	5	7	7	10	11	14	15
10	0	0	3	3	3	6	9	9	12	15	18	21	27	30	39	45
11	0	0	2	1	3	0	4	3	5	3	7	5	10	8	14	12
12	0	0	$\frac{3}{2}$	3	3	6	6	6	9	12	15	18	21	24	30	36
*13	0	0	0	−1	1	1	2	0	3	3	4	3	8	6	9	8
14	0	0	0	3	3	3	3	6	6	9	9	9	15	18	18	24
*15	0	0	0	0	2	−1	1	1	3	0	4	1	5	2	5	4
16	0	0	0	0	0	0	3	0	3	3	3	6	6	6	9	12
*17	0	0	0	0	1	0	2	0	2	−1	1	0	3	1	3	0

m = 25 $c\,=\,1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1
2	0	0	0	1	0	0	1	0	1	1	1	1	1	1	1	2
3	0	0	0	0	1	0	1	1	1	1	1	1	2	2	2	2
4	0	1	1	0	1	1	1	1	1	1	2	2	2	3	3	3
5	0	0	1	1	0	1	1	1	1	2	2	2	3	3	3	4
6	0	0	0	0	1	1	1	1	2	2	2	2	3	3	4	4
7	0	1	0	1	1	1	1	2	2	2	2	3	3	4	4	5
8	0	0	1	1	1	1	1	1	2	2	3	3	4	3	5	5
9	0	1	1	1	1	1	2	2	2	2	3	3	4	4	5	6
10	0	$\frac{1}{2}$	1	0	1	1	2	2	2	2	3	3	4	4	5	5
11	0	0	0	1	1	1	1	1	2	3	2	3	4	4	5	6
12	0	0	1	1	1	1	2	1	2	2	3	3	4	4	5	5
13	0	0	1	0	1	1	1	2	2	2	3	3	3	4	5	5
14	0	0	1	1	1	1	2	1	2	2	3	3	4	3	5	5
15	0	0	0	1	1	1	1	1	2	2	2	2	3	4	4	5
16	0	0	0	0	1	0	1	1	2	1	2	2	3	3	4	4
17	0	0	0	1	0	1	1	1	1	2	2	2	3	3	3	4
18	0	0	0	0	1	0	1	1	1	1	2	1	3	2	3	3
19	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	3
20	0	0	0	0	$\frac{1}{2}$	0	1	0	1	1	1	1	2	1	3	2
21	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2
22	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	1
23	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1
24	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0

A.3 Table of $\mathcal{Q}_M$ (weight 1 case)

We give the coefficients of $\mathcal{Q}_M$ for all products of three primes less than 150 for which $\mathcal{Q}_M$. We also show the factor c_M relating $\mathcal{Q}_M$ to $\mathcal{F}_M$ and the minimal discriminant of $\mathcal{F}_M$, since it is no longer assumed to be -1. This time an asterisk means simply that the corresponding row contains (non-polar) coefficients of opposite signs, since here there are many examples of values of $r \pmod{2m}$ for which the scalar factor κ_r that determines the sign of the coefficients $C(\Delta, r)$ for large Δ is negative rather than positive for $r = r_{\min}$.

M = 30 $c_M = 3$ $\Delta_{\min} = -1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	1	1	2	1	3	2	3	3	5	3	6	5	7	7	9
7	0	1	2	2	3	3	4	4	6	5	7	8	9	9	12	12

M = 42 $c_M = 2$ $\Delta_{\min} = -1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	−1	0	−1	−1	−1	0	−2	−1	−2	−1	−2	−1	−3	−2	−3
5	0	1	1	1	2	1	2	2	2	3	3	2	4	4	4	4
11	0	1	1	2	1	2	2	3	2	3	3	4	4	5	4	6

M = 66 $c_M = 12$ $\Delta_{\min} = -25$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10
1	−5	−5	−15	−60	−125	−313	−620	−1270	−2358	−4394	−7698
5	−1	15	65	175	450	989	2105	4140	7930	14508	25915
7	0	16	55	155	385	852	1816	3597	6880	12645	22627
13	0	5	22	60	155	357	781	1567	3070	5725	10367
19	0	0	−5	−33	−99	−268	−605	−1320	−2623	−5104	−9398

M = 70 $c_M = 1$ $\Delta_{\min} = -1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	0	−1	−1	0	0	−1	0	−1	−1	−1	−1	−1	0	−1	−1
3	0	0	1	0	1	0	0	1	1	0	1	1	1	0	1	1
9	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1	2
11	0	−1	0	0	0	−1	0	−1	0	0	0	−1	0	−1	0	−1
13	0	1	1	0	1	1	1	1	1	1	1	1	1	2	2	1
23	0	0	1	1	1	1	1	1	1	1	2	1	2	1	2	2

M = 78 $c_M = 1$ $\Delta_{\min} = -1$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	−1	0	0	−1	−1	0	0	−1	0	0	−1	0	−1	−1	−1	−1
5	0	1	0	1	0	1	0	1	1	1	0	1	1	1	1	1
7	0	−1	0	0	0	0	0	−1	0	−1	0	0	0	−1	0	0
11	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1
17	0	1	0	1	1	1	1	1	0	1	1	1	1	2	1	2
23	0	0	1	1	1	0	1	1	1	1	1	1	2	1	1	1

M = 102 $c_M = 6$ $\Delta_{\min} = -25$

n / $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11
1	−3	−1	−8	−14	−31	−56	−102	−173	−293	−461	−732	−1129
5	−1	7	16	41	76	153	262	454	745	1215	1886	2941
7	0	3	12	21	45	76	146	236	408	636	1024	1555
11	0	3	13	26	59	103	197	327	560	890	1427	2187
13	0	6	15	31	68	125	229	389	657	1055	1688	2603
19	0	2	7	16	35	61	120	205	349	574	924	1433
*25	0	0	1	−1	−4	−11	−16	−42	−61	−115	−178	−303
31	0	0	0	−3	−12	−27	−61	−113	−209	−358	−605	−972

M = 105 $c_M = 1$ $\Delta_{\min} = -4$

n / $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	1	1	2	1	2	3	3	3	4	4	5	6	7	7
2	−1	−1	−1	−1	−1	−1	−2	−2	−2	−3	−3	−3	−4	−4	−5	−6
4	0	1	0	1	1	1	1	1	2	2	2	2	3	4	3	4
8	0	0	1	1	0	1	2	1	2	2	2	3	3	3	4	5
11	0	0	1	0	1	1	1	1	2	1	3	2	3	2	4	4
13	0	1	0	1	1	1	1	2	1	2	2	3	3	4	3	5
16	0	1	1	1	2	2	2	3	3	4	4	5	6	7	8	8
17	0	−1	0	−1	0	−1	0	−2	0	−1	−1	−2	−1	−3	−1	−3
19	0	1	1	2	1	2	2	3	3	4	4	5	5	7	7	9
23	0	0	1	0	1	1	2	1	2	1	3	3	4	3	5	4
32	0	0	0	0	0	1	0	0	1	1	1	1	1	1	2	2
34	0	0	0	1	1	1	1	2	2	2	2	3	4	4	4	6

M = 110 $c_M = 3$ $\Delta_{\min} = -9$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−2	−1	−2	−3	−3	−5	−5	−9	−9	−13	−15	−22	−25	−32	−37
3	−1	0	2	1	4	4	6	7	11	14	18	21	28	34	44
7	0	1	1	3	2	3	4	5	6	9	9	13	16	20	21
9	0	2	3	2	5	5	9	10	13	15	22	24	33	39	50
13	0	1	2	4	6	7	10	14	18	23	31	37	49	58	74
17	0	2	3	4	5	8	10	12	17	22	27	34	43	52	65
19	0	1	2	2	2	3	5	5	6	9	10	13	16	19	24
23	0	0	2	1	4	4	6	8	13	13	20	23	31	37	49
$\ast$29	0	0	1	0	1	0	1	−2	0	−3	−2	−4	−2	−9	−6
$\ast$39	0	0	0	0	1	−1	−1	−2	−3	−6	−5	−10	−11	−16	−20

M = 114 $c_M = 4$ $\Delta_{\min} = -25$

n $r_{\min}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	−1	4	9	18	37	62	110	181	291	449	695	1034	1537	2235
5	−1	0	4	3	11	16	32	41	83	113	188	266	413	574
7	0	6	12	30	53	99	166	282	444	706	1067	1622	2387	3498
11	0	3	11	21	42	74	133	216	351	547	849	1271	1897	2757
13	0	3	8	16	37	59	111	178	295	457	718	1062	1606	2327
17	0	2	9	19	38	70	127	205	345	538	840	1267	1904	2778
23	0	0	3	5	15	24	48	75	135	204	333	496	765	1110
29	0	0	−1	−2	−6	−10	−20	−37	−64	−100	−165	−257	−388	−583
35	0	0	0	−3	−7	−18	−32	−67	−111	−194	−310	−501	−767	−1184

M = 130 $c_M = 2$ $\Delta_{\min} = -9$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	−1	0	−1	0	−1	−1	−2	−2	−3	−3	−5	−5	−7	−7	−12
3	−1	−1	−1	−2	−1	−3	−3	−5	−5	−7	−7	−11	−11	−16	−17
7	0	1	2	2	3	3	6	6	8	9	13	14	19	21	27
9	0	1	2	2	4	4	6	8	10	12	16	18	24	28	35
11	0	1	1	2	2	3	4	4	6	8	9	10	14	17	19
17	0	1	1	2	2	3	4	5	5	8	9	12	14	17	19
19	0	1	2	2	4	5	6	8	11	13	17	20	26	31	38
21	0	1	2	3	3	5	6	8	10	13	15	20	23	29	35
27	0	0	1	0	1	1	2	0	3	1	3	3	5	3	7
29	0	0	1	1	2	2	4	4	6	6	10	11	15	16	22
37	0	0	0	0	0	−1	0	−1	−2	−2	−2	−4	−4	−6	−7
47	0	0	0	0	0	−1	0	−2	−2	−3	−3	−6	−5	−9	−10

M = 138 $c_M = 2$ $\Delta_{\min} = -49$

$r_{\min}$ \ n	0	1	2	3	4	5	6	7	8	9	10	11
1	−5	5	20	57	141	323	658	1308	2449	4450	7786	13373
5	−2	9	27	56	125	235	451	812	1442	2463	4182	6884
7	−1	13	44	111	265	555	1138	2170	4032	7183	12541	21237
11	0	14	46	115	261	546	1091	2055	3787	6702	11610	19591
13	0	9	29	68	151	308	612	1139	2080	3675	6323	10653
17	0	6	28	68	174	354	753	1419	2678	4775	8407	14255
∗19	0	2	6	−2	−12	−69	−157	−397	−784	−1581	−2884	−5240
25	0	0	−3	−23	−69	−192	−427	−926	−1817	−3473	−6310	−11189
31	0	0	−1	−11	−33	−100	−220	−491	−973	−1890	−3456	−6203
∗37	0	0	0	−1	−1	−5	6	8	49	110	263	490
43	0	0	0	0	2	13	47	128	309	677	1369	2653

B. Bibliography

[1] Sergei Alexandrov, Jan Manschot, and Boris Pioline, *D3-instantons, Mock Theta Series and Twistors*, (2012), `arxiv:1207.1109`.

[2] Murad Alim, Babak Haghighat, Michael Hecht, Albrecht Klemm, Marco Rauch, and Thomas Wotschke, *Wall-crossing holomorphic anomaly and mock modularity of multiple M5-branes*, (2010), `arxiv:1012.1608`.

[3] O. Alvarez, T. P. Killingback, M. L. Mangano, and P. Windey, *String theory and loop space index theorems*, Commun. Math. Phys. **111** (1987) 1.

[4] S. K. Ashok and J. Troost, *A Twisted Non-compact Elliptic Genus*, JHEP **1103** (2011) 067, `arxiv:1101.10593`.

[5] Shamik Banerjee and Ashoke Sen, *Duality orbits, dyon spectrum and gauge theory limit of heterotic string theory on T^6*, JHEP **0803** (2008) 022, `arxiv:0712.0043`.

[6] ______, *S-duality action on discrete T-duality invariants*, JHEP **0804** (2008) 012, `arxiv:0801.0149`.

[7] S. Banerjee, A. Sen, and Y. K. Srivastava, *Generalities of quarter BPS dyon partition function and dyons of torsion two*, JHEP **0805** (2008) 101, `arxiv:0802.0544`.

[8] ______, *Partition functions of torsion > 1 dyons in heterotic string theory on T^6*, JHEP **0805** (2008) 098, `arxiv:0802.1556`.

[9] ______, *Genus two surface and quarter BPS dyons: The contour prescription*, JHEP **0903** (2009) 151, `arxiv:0808.1746`.

[10] N. Banerjee, I. Mandal and A. Sen, *Black hole hair removal*, JHEP **0907** (2009) 091, `arxiv:0901.0359`.

[11] B. Bates and F. Denef, *Exact solutions for supersymmetric stationary black hole composites*, JHEP **1111** (2011) 127, `hep-th/0304094`.

[12] J. C. Breckenridge, Robert C. Myers, A. W. Peet, and C. Vafa, *D-branes and spinning black holes*, Phys.Lett. **B391** (1997) 93–98, `hep-th/9602065`.

[13] K. Bringmann and K. Ono, *Coefficients of harmonic Maass forms*, Proceedings of the 2008 University of Florida Conference on Partitions, q-series, and modular forms.

[14] K. Bringmann and O. Richter, *Zagier-type dualities and lifting maps for harmonic Maass-Jacobi forms*, Adv. Math. 225 **no. 4,** (2010), 2298–2315.

[15] K. Bringmann and J. Manschot, *From sheaves on P^2 to a generalization of the Rademacher expansion*, Amer. J. Math. **135** (2013), no. 4, 1039-1065. `arXiv:1006.0915`.

[16] K. Bringmann and K. Mahlburg, *An extension of the Hardy-Ramanujan Circle Method and applications to partitions without sequences*, American Journal of Math 133 (2011), pages 1151-1178.

[17] K. Bringmann and A. Folsom, *Almost harmonic Maass forms and Kac-Wakimoto characters*, Journal für die Reine und Angewandte Mathematik, doi:10.1515/crelle-2012-0102, `arXiv:1112.4726`.

[18] K. Bringmann, M. Raum, and O. Richter, *Harmonic Maass-Jacobi forms with singularities and a theta-like decomposition*, to appear in Transactions AMS, `arXiv:1207.5600`.

[19] K. Bringmann and S. Murthy, *On the positivity of black hole degeneracies in string theory*, Commun. Number Theory Phys. **7** (2013), no. 1, 15-56. `arXiv:1208.3476`.

[20] K. Bringmann and A. Folsom, *On the asymptotic behavior of Kac-Wakimoto characters*, Proceedings of the American Mathematical Society 141 (2013), pages 1567–1576.

[21] J. H. Bruinier, G. van der Geer, G. Harder and D. Zagier, *The* 1-2-3 *of Modular Forms*, Universitext, Springer-Verlag, Belin–Heidelberg (2008).

[22] G. L. Cardoso, J. R. David, B. de Wit, and S. Mahapatra, *The mixed black hole partition function for the STU model*, JHEP **12** (2008), 086, `arxiv:0810.1233`.

[23] Alejandra Castro and Sameer Murthy, *Corrections to the statistical entropy of five dimensional black holes*, JHEP **0906** (2009), 024, `arxiv:0807.0237`.

[24] Miranda C.N. Cheng, *K3 surfaces, $N = 4$ Dyons, and the Mathieu group M24*, (2010), `arxiv:1005.5415`.

[25] Miranda C. N. Cheng, John F. R. Duncan and Jeffrey A. Harvey, *Umbral Moonshine*, `arXiv:1204.2779`.

[26] Miranda C. N. Cheng and Erik Verlinde, *Dying dyons don't count*, JHEP **0709** (2007) 070, `arxiv:0706.2363`.

[27] Young Ju Choie and Subong Lim, *The heat operator and mock Jacobi forms*, The Ramanujan Journal Volume 22, Number 2 (2010), 209-219.

[28] Atish Dabholkar, *Exact counting of black hole microstates*, Phys. Rev. Lett. **94** (2005), 241301, `hep-th/0409148`.

[29] Atish Dabholkar, Frederik Denef, Gregory W. Moore, and Boris Pioline, *Exact and asymptotic degeneracies of small black holes*, JHEP **08** (2005), 021, `hep-th/0502157`.

[30] ______, *Precision counting of small black holes,* JHEP **0510** (2005) 096, `hep-th/0507014`.

[31] Atish Dabholkar, Davide Gaiotto, and Suresh Nampuri, *Comments on the spectrum of CHL dyons*, JHEP **01** (2008), 023, `hep-th/0702150`.

[32] Atish Dabholkar, Gary W. Gibbons, Jeffrey A. Harvey, and Fernando Ruiz Ruiz, *Superstrings and solitons*, Nucl. Phys. **B340** (1990), 33–55.

[33] Atish Dabholkar, Joao Gomes, and Sameer Murthy, *Counting all dyons in $N = 4$ string theory*, JHEP **1105** (2011) 059 `arxiv:0803.2692`.

[34] ———, *Quantum black holes, localization and the topological string*, JHEP **1106** (2011) 019 `arxiv:1012.0265`.

[35] ———, *Localization & Exact Holography*, JHEP **1304** (2013) 062 `arxiv:1111.1161`.

[36] ———, *Nonperturbative black hole entropy and Kloosterman sums*, (2014), `arxiv:1404.0033`.

[37] Atish Dabholkar, Joao Gomes, Sameer Murthy, and Ashoke Sen, *Supersymmetric index from black hole entropy*, (2010), JHEP **1104** (2011) 034 `arxiv:1009.3226`.

[38] A. Dabholkar, M. Guica, S. Murthy and S. Nampuri, *No entropy enigmas for $N = 4$ dyons*, JHEP **1006** (2010) 007 `arxiv:0903.2481`.

[39] Atish Dabholkar and Jeffrey A. Harvey, *Nonrenormalization of the superstring tension*, Phys. Rev. Lett. **63** (1989), 478.

[40] Atish Dabholkar, Renata Kallosh, and Alexander Maloney, *A stringy cloak for a classical singularity*, JHEP **12** (2004), 059, `hep-th/0410076`.

[41] Atish Dabholkar and Suresh Nampuri, *Spectrum of dyons and black Holes in CHL orbifolds using Borcherds lift*, JHEP **11** (2007), 077, `hep-th/0603066`.

[42] Justin R. David and Ashoke Sen, *CHL dyons and statistical entropy function from D1-D5 system*, JHEP **11** (2006), 072, `hep-th/0605210`.

[43] L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, *Strings on Orbifolds*, Nucl. Phys. B **261** (1985) 678.

[44] Jan de Boer, Miranda C. N. Cheng, Robbert Dijkgraaf, Jan Manschot, and Erik Verlinde, *A Farey tail for attractor black holes*, JHEP **11** (2006), 024, `hep-th/0608059`.

[45] J. de Boer, F. Denef, S. El-Showk, I. Messamah and D. Van den Bleeken, *Black hole bound states in $AdS_3 \times S^2$*, JHEP **0811** (2008) 050, `arXiv:0802.2257`.

[46] F. Denef, *Supergravity flows and D-brane stability*, JHEP **0008** (2000) 050, `hep-th/0005049`.

[47] F. Denef, *Quantum quivers and Hall/hole halos*, JHEP **0210** (2002) 023, `hep-th/0206072`.

[48] Frederik Denef and Gregory W. Moore, *Split states, entropy enigmas, holes and halos*, (2007), `hep-th/0702146`.

[49] Robbert Dijkgraaf, Juan Martin Maldacena, Gregory Winthrop Moore, and Erik P. Verlinde, *A black hole Farey tail*, (2000), `hep-th/0005003`.

[50] Robbert Dijkgraaf, Gregory W. Moore, Erik P. Verlinde, and Herman L. Verlinde, *Elliptic genera of symmetric products and second quantized strings*, Commun. Math. Phys. **185** (1997), 197–209, `hep-th/9608096`.

[51] Robbert Dijkgraaf, Erik P. Verlinde, and Herman L. Verlinde, *Counting dyons in* $N = 4$ *string theory*, Nucl. Phys. **B484** (1997), 543–561, `hep-th/9607026`.

[52] Tohru Eguchi, Hirosi Ooguri, and Yuji Tachikawa, *Notes on the* $K3$ *surface and the Mathieu group* M_{24}, (2010), `arxiv:1004.0956`.

[53] Tohru Eguchi, Hirosi Ooguri, S-K. Yang and Anne Taormina, *Superconformal algebras and string compactification on manifolds with* $SU(N)$ *Holonomy*, Nucl. Phys. **B315** (1989), 193.

[54] Tohru Eguchi and Yuji Sugawara, $SL(2,R)/U(1)$ *supercoset and elliptic genera of noncompact Calabi-Yau manifolds*, JHEP **0405** (2004) 014. `hep-th/0403193`

[55] Tohru Eguchi, Yuji Sugawara and Anne Taormina, *Liouville field, modular forms and elliptic genera*, JHEP **0703** (2007), 119. `hep-th/0611338`

[56] Tohru Eguchi and Anne Taormina, *On the unitary representation of* $N = 2$ *and* $N = 4$ *superconformal algebras*, Phys.Lett. **B210** (1988),125.

[57] M. Eichler and D. Zagier, *The Theory of Jacobi Forms*, Birkhäuser, 1985.

[58] Sergio Ferrara, Renata Kallosh, and Andrew Strominger, $N = 2$ *extremal black holes*, Phys. Rev. **D52** (1995), 5412–5416, `hep-th/9508072`.

[59] Davide Gaiotto, *Re-recounting dyons in* $N = 4$ *string theory*, (2005), `hep-th/0506249`.

[60] Davide Gaiotto, Andrew Strominger, and Xi Yin, *New connections between 4D and 5D black holes*, JHEP **02** (2006), 024, `hep-th/0503217`.

[61] ______, *The M5-brane elliptic genus: Modularity and BPS states*, JHEP **08** (2007), 070, `hep-th/0607010`.

[62] Jerome P. Gauntlett, Nakwoo Kim, Jaemo Park, and Piljin Yi, *Monopole dynamics and BPS dyons in* $N = 2$ *super-Yang-Mills theories*, Phys. Rev. **D61** (2000), 125012, `hep-th/9912082`.

[63] Paul H. Ginsparg, *Applied conformal field theory*, (1988), `hep-th/9108028`.

[64] Amit Giveon and David Kutasov, *Little string theory in a double scaling limit*, JHEP **9910** (1999) 034, `hep-th/9909110`.

[65] L. Goettsche, *The Betti numbers of the Hilbert scheme of points on a smooth protective surface*, Math. Ann. **286** (1990), 193.

[66] V. Gritsenko, "Elliptic genus of Calabi-Yau manifolds and Jacobi and Siegel modular forms," `math/9906190`

[67] V. A. Gritsenko and V. V. Nikulin, *Automorphic forms and Lorentzian Kac–Moody Algebras. Part I*, (1996) `alg-geom/9610022`.

[68] V. A. Gritsenko and V. V. Nikulin, *Automorphic forms and Lorentzian Kac–Moody Algebras. Part II*, (1996) `alg-geom/9611028`.

[69] V. A. Gritsenko, N.-P. Skoruppa, D. Zagier, *Theta blocks I: Elementary theory and examples*, in preparation.

[70] Rajesh Kumar Gupta and Sameer Murthy, *All solutions of the localization equations for N=2 quantum black hole entropy*, JHEP **1302** (2013) 141, `arXiv:1208.6221`.

[71] Jeffrey A. Harvey and Sameer Murthy, *Moonshine in Fivebrane Spacetimes*, JHEP **1401** (2014) 146, `arXiv:1307.7717`.

[72] D. R. Hickerson, *On the seventh order mock theta functions*, Inv. Math. **94**, (1988), 661–677.

[73] F. Hirzebruch and D. Zagier, *Intersection numbers on curves on Hilbert modular surfaces and modular forms of Nebentypus*, Inv. Math. **36** (1976), 57.

[74] C. M. Hull and P. K. Townsend, *Unity of superstring dualities*, Nucl. Phys. **B438** (1995), 109–137, `hep-th/9410167`.

[75] D. P. Jatkar, A. Sen and Y. K. Srivastava, *Black hole hair removal: non-linear analysis*, JHEP **1002** (2010) 038, `arXiv:0907.0593`.

[76] D. Joyce, Y. Song, *A Theory of Generalized Donaldson-Thomas Invariants*, Memoirs of the AMS 217 (2012), `arXiv:0810.5645`.

[77] Toshiya Kawai, Yasuhiko Yamada, and Sung-Kil Yang, *Elliptic genera and* $N = 2$ *superconformal field theory*, Nucl. Phys. **B414** (1994), 191–212, `hep-th/9306096`.

[78] M. Kontsevich and Y. Soibelman, *Stability structures, motivic Donaldson-Thomas invariants and cluster transformations*, `arXiv:0811.2435`.

[79] Per Kraus and Finn Larsen, *Microscopic black hole entropy in theories with higher derivatives*, JHEP **09** (2005), 034, `hep-th/0506176`.

[80] ———, *Partition functions and elliptic genera from supergravity*, (2006), hep-th/0607138.

[81] Gabriel Lopes Cardoso, Bernard de Wit, and Thomas Mohaupt, *Corrections to macroscopic supersymmetric black-hole entropy*, Phys. Lett. **B451** (1999), 309–316, hep-th/0412287.

[82] ———, *Area law corrections from state counting and supergravity*, Class. Quant. Grav. **17** (2000), 1007–1015, hep-th/9910179.

[83] ———, *Deviations from the area law for supersymmetric black holes*, Fortsch. Phys. **48** (2000), 49–64, hep-th/9904005.

[84] ———, *Macroscopic entropy formulae and non-holomorphic corrections for supersymmetric black holes*, Nucl. Phys. **B567** (2000), 87–110, hep-th/9906094.

[85] Gabriel Lopes Cardoso, B. de Wit, J. Kappeli, and T. Mohaupt, *Asymptotic degeneracy of dyonic $N = 4$ string states and black hole entropy*, JHEP **12** (2004), 075.

[86] J. M. Maldacena and A. Strominger, *AdS(3) black holes and a stringy exclusion principle*, JHEP **9812** (1998) 005 hep-th/9804085.

[87] Juan M. Maldacena, Andrew Strominger, and Edward Witten, *Black hole entropy in M-theory*, JHEP **12** (1997), 002, hep-th/9711053.

[88] J. Manschot, B. Pioline and A. Sen, *Wall Crossing from Boltzmann Black Hole Halos*, JHEP **1107** (2011) 059 arXiv:1011.1258.

[89] Jan Manschot and Gregory W. Moore, *A Modern Farey Tail*, (2007), arxiv:0712.0573.

[90] Jan Manschot, *Stability and duality in N=2 supergravity*, Commun.Math.Phys., **299**, 651-676 (2010), arxiv:0906.1767.

[91] J. A. Minahan, D. Nemeschansky, C. Vafa, and N. P. Warner, *E-strings and $N = 4$ topological Yang-Mills theories*, Nucl. Phys. **B527** (1998), 581–623, hep-th/9802168.

[92] Ruben Minasian, Gregory W. Moore, and Dimitrios Tsimpis, *Calabi-Yau black holes and $(0, 4)$ sigma models*, Commun. Math. Phys. **209** (2000), 325–352, hep-th/9904217.

[93] Gregory W. Moore, *Les Houches lectures on strings and arithmetic*, (2004), hep-th/0401049.

[94] Sameer Murthy, *Notes on noncritical superstrings in various dimensions*, JHEP **0311** (2003), 056, hep-th/0305197.

[95] Sameer Murthy and Boris Pioline, *A Farey tale for $N = 4$ dyons*, JHEP **0909** (2009), 022, arxiv:0904.4253.

[96] Sameer Murthy and Valentin Reys, *Quantum black hole entropy and the holomorphic prepotential of N=2 supergravity*, JHEP **10** (2013) 099, arXiv:1306.3796.

[97] Serge Ochanine, *Sur les genres multiplicatifs definis par des integrales elliptiques*, Topology **26** (1987), 143.

[98] René Olivetto, *On the Fourier coefficients of meromorphic Jacobi forms*, `arXiv:1210.7926`.

[99] H. Ooguri, *Superconformal symmetry and geometry of Ricci flat Kähler manifolds*, Int. J. Mod. Phys. **A4** (1989), 4303–4324.

[100] C. N. Pope, *Axial vector anomalies and the index theorem in charged Schwarzschild and Taub - NUT spaces*, Nucl. Phys. **B141** (1978), 432.

[101] Ashoke Sen, *Dyon-monopole bound states, selfdual harmonic forms on the multi-monopole moduli space, and* $Sl(2,\mathbb{Z})$ *invariance in string theory*, Phys. Lett. **B329** (1994), 217–221, `hep-th/9402032`.

[102] ______, *Strong-weak coupling duality in four-dimensional string theory*, Int. J. Mod. Phys. **A9** (1994), 3707–3750, `hep-th/9402002`.

[103] ______, *Extremal black holes and elementary string states*, Mod. Phys. Lett. **A10** (1995), 2081–2094, `hep-th/9504147`.

[104] ______, *Black holes and the spectrum of half-BPS states in* $N = 4$ *supersymmetric string theory*, (2005), `hep-th/0504005`.

[105] ______, *Walls of marginal stability and dyon spectrum in* $N = 4$ *supersymmetric string theories*, JHEP **05** (2007), 039, `hep-th/0702141`.

[106] ______, *Entropy Function and* $AdS(2)/CFT(1)$ *Correspondence*, JHEP **11** (2008), 075, `arxiv:0805.0095`.

[107] ______, *Quantum Entropy Function from* $AdS(2)/CFT(1)$ *Correspondence*, (2008), `arxiv:0809.3304`.

[108] ______, *Negative discriminant states in* $N = 4$ *supersymmetric string theories*, (2011), `arxiv:1104.1498`.

[109] David Shih, Andrew Strominger, and Xi Yin, *Recounting dyons in* $N = 4$ *string theory*, JHEP **10** (2006), 087, `hep-th/0505094`.

[110] David Shih and Xi Yin, *Exact black hole degeneracies and the topological string*, JHEP **04** (2006), 034, `hep-th/0508174`.

[111] N.-P. Skoruppa, *Über den Zusammenhang zwischen Jacobi-Formen und Modulformen halbganzen Gewichts*, Dissertation, Universität Bonn (1984).

[112] ______, *Developments in the theory of Jacobi forms*. In Automorphic functions and their applications Khabarovsk, 1988, Acad. Sci. USSR Inst. Appl. Math., Khabarovsk 1990, 167–185

[113] ———, *Binary quadratic forms and the Fourier coefficients of elliptic and Jacobi modular forms.* J. reine Angew. Math. 411 (1990), 66–95

[114] N.-P. Skoruppa, D. Zagier, *A trace formula for Jacobi forms* J. reine Angew. Math. 393 (1989) 168-198.

[115] ———, *Jacobi forms and a certain space of modular forms.* Inventiones mathematicae 94, 113–146 (1988).

[116] Andrew Strominger, *Macroscopic entropy of* $N = 2$ *extremal black holes*, Phys. Lett. **B383** (1996), 39–43, `hep-th/9602111`.

[117] ———, $AdS(2)$ *quantum gravity and string theory*, JHEP **01** (1999), 007, `hep-th/9809027`.

[118] Andrew Strominger and Cumrun Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, Phys. Lett. **B379** (1996), 99–104, `hep-th/9601029`.

[119] Jan Troost, *The non-compact elliptic genus: mock or modular*, JHEP **1006** (2010), 104, `arxiv:1004.3649`.

[120] Cumrun Vafa and Edward Witten, *A strong coupling test of S-duality*, Nucl. Phys. **B431** (1994), 3–77, `hep-th/9408074`.

[121] Katrin Wendland, Ph.D. Thesis, Bonn University (2000).

[122] Edward Witten, *Dyons of charge* $e\theta/2\pi$, Phys. Lett. B **86** (1979) 283.

[123] ———, *Elliptic genera and quantum field theory*, Commun. Math. Phys. **109** (1987), 525.

[124] ———, *String theory dynamics in various dimensions*, Nucl. Phys. **B443** (1995), 85–126, `hep-th/9503124`.

[125] Edward Witten and David I. Olive, *Supersymmetry algebras that include topological charges*, Phys. Lett. **B78** (1978), 97.

[126] D. Zagier, *Nombres de classes et formes modulaires de poids* 3/2, C. R. Acad. Sc. Paris **281** (1975), 883.

[127] ———, *The Rankin-Selberg method for automorphic functions which are not of rapid decay*, J. Fac. Sci. Tokyo 28 (1982) 415-438.

[128] ———, *Traces of singular moduli*, in “Motives, Polylogarithms and Hodge Theory” (eds. F. Bogomolov, L. Katzarkov), Lecture Series **3**, International Press, Somerville (2002), 209–244.

[129] ———, *Ramanujan’s mock theta functions and their applications* [*d’après Zwegers and Bringmann-Ono*], Séminaire Bourbaki, 60ème année, 2006-2007, **986** (2007).

[130] S. P. Zwegers, *Mock theta functions*, Thesis, Utrecht (2002), `http://igitur-archive.library.uu.nl/dissertations/2003-0127-094324/inhoud.htm`.

Monstrous moonshine and monstrous Lie superalgebras.

Invent. Math. 109, 405-444 (1992).

Richard E. Borcherds,

Department of pure mathematics and mathematical statistics, 16 Mill Lane, Cambridge CB2 1SB, England.

We prove Conway and Norton's moonshine conjectures for the infinite dimensional representation of the monster simple group constructed by Frenkel, Lepowsky and Meurman. To do this we use the no-ghost theorem from string theory to construct a family of generalized Kac-Moody superalgebras of rank 2, which are closely related to the monster and several of the other sporadic simple groups. The denominator formulas of these superalgebras imply relations between the Thompson functions of elements of the monster (i.e. the traces of elements of the monster on Frenkel, Lepowsky, and Meurman's representation), which are the replication formulas conjectured by Conway and Norton. These replication formulas are strong enough to verify that the Thompson functions have most of the "moonshine" properties conjectured by Conway and Norton, and in particular they are modular functions of genus 0. We also construct a second family of Kac-Moody superalgebras related to elements of Conway's sporadic simple group Co_1. These superalgebras have even rank between 2 and 26; for example two of the Lie algebras we get have ranks 26 and 18, and one of the superalgebras has rank 10. The denominator formulas of these algebras give some new infinite product identities, in the same way that the denominator formulas of the affine Kac-Moody algebras give the Macdonald identities.

1 Introduction.
2 Introduction (continued).
3 Vertex algebras.
4 Generalized Kac-Moody algebras.
5 The no-ghost theorem.
6 Construction of the monster Lie algebra.
7 The simple roots of the monster Lie algebra.
8 The twisted denominator formula.
9 The moonshine conjectures.
10 The monstrous Lie superalgebras.
11 Some modular forms.
12 The fake monster Lie algebra.
13 The denominator formula for fake monster Lie algebras.
14 Examples of fake monster Lie algebras.
15 Open problems.

1 Introduction.

The main result of the first half of this paper is the following.

Theorem 1.1. *Suppose that $V=\oplus_{n\in\mathbf{Z}}V_n$ is the infinite dimensional graded representation of the monster simple group constructed by Frenkel, Lepowsky, and Meurman [16,17]. Then for any element g of the monster the Thompson series $T_g(q)=\sum_{n\in\mathbf{Z}}\mathrm{Tr}(g|V_n)q^n$ is* *a Hauptmodul for a genus 0 subgroup of $SL_2(\mathbf{R})$, i.e., V satisfies the main conjecture in Conway and Norton's paper [13].*

We prove this by constructing a $\mathbf{Z}^2$-graded Lie algebra acted on by the monster, called the monster Lie algebra. This is a generalized Kac-Moody algebra, and by calculating the "twisted denominator formulas" of this Lie algebra explicitly we get enough information about the Thompson series $T_g(q)$ to determine them.

In this introduction we explain this result in more detail and briefly describe the proof, which is contained in sections 6 to 9. The rest of this paper is organized as follows. Section 2 is an introduction to the second half of the paper (sections 10 to 14) which uses some of the techniques of the proof to find some new infinite product formulas and infinite dimensional Lie algebras. Sections 3, 4, and 5 summarize some known results about vertex algebras, Kac-Moody algebras, and the no-ghost theorem that we use in the proof of theorem 1.1. Section 15 contains a list of some open questions. There is a list of notation we use at the end of section 1.

The Fischer-Griess monster sporadic simple group, of order $2^{46}3^{20}5^97^611^213^317.19.23.29.31.41.47.59.71$, acts naturally and explicitly on a graded real vector space $V = \oplus_{n\in\mathbf{Z}} V_n$ constructed by Frenkel, Lepowsky and Meurman [16,17] (The vector spaces V and V_n are denoted by $V^\natural$ and $V^\natural_{-n}$ in [16].) The dimension of V_n is equal to the coefficient $c(n)$ of the elliptic modular function $j(\tau) - 744 = \sum_n c(n)q^n = q^{-1} + 196884q + 21493760q^2 + \dots$ (where we write q for $e^{2\pi i\tau}$, and $\mathrm{Im}(\tau) > 0$). One of the main remaining problems from [16], which theorem 1.1 solves, is to calculate the character of V as a graded representation of the monster, or in other words to calculate the trace $\mathrm{Tr}(g|V_n)$ of each element g of the monster on each space V_n. The best way to describe this information is to define the Thompson series

$$T_g(q) = \sum_{n\in\mathbf{Z}} \mathrm{Tr}(g|V_n)q^n$$

for each element g of the monster, so we want to calculate these Thompson series. For example, if 1 is the identity element of the monster then $\mathrm{Tr}(1|V_n) = \dim(V_n) = c(n)$, so that the Thompson series $T_1(q) = j(\tau) - 744$ is the elliptic modular function. McKay, Thompson, Conway and Norton conjectured [13] that the Thompson series $T_g(q)$ are all Hauptmoduls for certain explicitly given modular groups of genus 0. (More precisely, they only conjectured that there should be some graded module for the monster whose Thompson series are Hauptmoduls, since their conjectures came before the construction of V.) This conjecture follows from theorem 1.1. The corresponding Hauptmoduls are the ones listed in [13] (with their constant terms removed), so this completely describes V as a representation of the monster.

We recall the definition of a Hauptmodul. The group $SL_2(\mathbf{Z})$ acts on the upper half plane $H = \{\tau \in \mathbf{C}|\mathrm{Im}(\tau) > 0\}$ by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(\tau) = \frac{a\tau+b}{c\tau+d}$. A meromorphic function on H invariant under $SL_2(\mathbf{Z})$ and satisfying a certain regularity condition at $i\infty$ is called a modular function of level 1. The phrase "level 1" refers to the group $SL_2(\mathbf{Z})$; if this is replaced by some commensurable group we get modular functions of higher levels. The elliptic modular function $j(\tau)$ is, up to normalizations, the simplest nonconstant modular function of level 1; more precisely, the modular functions of level 1 are the rational functions of j. The element $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ of $SL_2(\mathbf{Z})$ takes τ to $\tau+1$, so in particular $j(\tau)$ is periodic and can be written as a Laurent series in $q=e^{2\pi i\tau}$. An exact expression for j is

$$j(\tau)=\frac{(1+240\sum_{n>0}\sigma_3(n)q^n)^3}{q\prod_{n>0}(1-q^n)^{24}}$$

where $\sigma_3(n)=\sum_{d|n}d^3$ is the sum of the cubes of the divisors of n; see any book on modular forms or elliptic functions, for example [30]. Another way of thinking about j is that it is an isomorphism from the quotient space $H/SL_2(\mathbf{Z})$ to the complex plane, which can be thought of as the Riemann sphere minus the point at infinity.

We can also consider functions invariant under some group G commensurable with $SL_2(\mathbf{Z})$ acting on H. The quotient H/G is again a compact Riemann surface $\overline{H/G}$ with a finite number of points removed. If this compact Riemann surface is a sphere, rather than something of higher genus, then we say that G is a genus 0 group. In this case a function giving an isomorphism from the compact Riemann surface $\overline{H/G}$ to the sphere $\mathbf{C}\cup\infty$ taking $i\infty$ to ∞ is called a Hauptmodul for the genus 0 group G; it is unique up to addition of a constant and multiplication by a nonzero constant. If a Hauptmodul can be written as $e^{-2\pi i a\tau}+$ a function vanishing at $i\infty$ for some positive a then we say that it is normalized. Every genus 0 group has a unique normalized Hauptmodul. For example, $j(\tau)-744=e^{-2\pi i\tau}+0+196884e^{2\pi i\tau}+\dots$ is the normalized Hauptmodul for the genus 0 group $SL_2(\mathbf{Z})$.

Another example is $G=\Gamma_0(2)$, where $\Gamma_0(N)=\{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in SL_2(\mathbf{Z})|c\equiv 0 \bmod N\}$. The quotient H/G is then a sphere with 2 points removed, so that G is a genus 0 group. Its normalized Hauptmodul is $T_{2-}(q)=24+q^{-1}\prod_{n>0}(1-q^{2n+1})^{24}=q^{-1}+276q-2048q^2+\dots$, and is equal to the Thompson series of a certain element of the monster of order 2 (of type 2B in atlas [14] notation). Similarly $\Gamma_0(N)$ is a genus 0 subgroup for several other values of N which correspond to elements of the monster. (However the genus of $\Gamma_0(N)$ tends to infinity as N increases, so there are only a finite number of integers N for which it has genus 0; more generally Thompson [31] has shown that there are only a finite number of conjugacy classes of genus 0 subgroups of $SL_2(\mathbf{R})$ which are commensurable with $SL_2(\mathbf{Z})$.)

So we want to calculate the Thompson series $T_g(\tau)$ and show that they are Hauptmoduls of genus 0 subgroups of $SL_2(\mathbf{R})$. The difficulty with doing this is as follows. Frenkel, Lepowsky, and Meurman constructed V as the sum of two subspaces V^+ and V^-, which are the $+1$ and -1 eigenspaces of a certain element of order 2 in the monster. If an element g of the monster commutes with this element of order 2, then it is not difficult to to work out its Thompson series $T_g(q)=\sum_n \mathrm{Tr}(g|V_n)q^n$ as the sum of two series given by its traces on V^+ and V^- (this is done in [16,17]), and it would probably be tedious but straightforward to check directly that these are all Hauptmoduls. Unfortunately, if an element of the monster is not conjugate to something that commutes with this element of order 2 then there is no obvious direct way of working out its Thompson series, because it muddles up V^+ and V^- in a very complicated way.

We calculate these Thompson series indirectly using the monster Lie algebra M. This is a $\mathbf{Z}^2=\mathbf{Z}\oplus\mathbf{Z}$ graded Lie algebra, whose piece of degree $(m,n)\in\mathbf{Z}^2$ is isomorphic as a module over the monster to V_{mn} if $(m,n)\neq(0,0)$ and to $\mathbf{R}^2$ if $(m,n)=(0,0)$, so for small degrees it looks like

$$\begin{array}{ccccccccc}
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & 0 & 0 & 0 & 0 & V_3 & V_6 & V_9 & \cdots \\
\cdots & 0 & 0 & 0 & 0 & V_2 & V_4 & V_6 & \cdots \\
\cdots & 0 & 0 & V_{-1} & 0 & V_1 & V_2 & V_3 & \cdots \\
\cdots & 0 & 0 & 0 & \mathbf{R}^2 & 0 & 0 & 0 & \cdots. \\
\cdots & V_3 & V_2 & V_1 & 0 & V_{-1} & 0 & 0 & \cdots \\
\cdots & V_6 & V_4 & V_2 & 0 & 0 & 0 & 0 & \cdots \\
\cdots & V_9 & V_6 & V_3 & 0 & 0 & 0 & 0 & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

Very briefly, this Lie algebra is constructed as the space of physical states of a bosonic string moving in a Z_2-orbifold of a 26-dimensional torus (or strictly speaking, about half the physical states). See section 6 for more details. The space of physical states is a subquotient of a vertex algebra constructed from the vertex algebra V; vertex algebras are described in more detail in [3,16,18], and the properties we use are summarized in section 3. This subquotient can be identified using the no-ghost theorem from string theory ([21] or section 5), and is as described above.

We need to know what the structure of the monster Lie algebra is. It turns out to be something called a generalized Kac-Moody algebra, so we explain what these are.

Section 4 describes the results about generalized Kac-Moody algebras that we use. This paragraph gives a brief summary of them. Kac-Moody algebras can be thought of as Lie algebras generated by a copy of sl_2 for each point in their Dynkin diagram. Generalized Kac-Moody algebras are rather like Kac-Moody algebras except that we are allowed to glue together the sl_2's in more complicated ways, and are also allowed to use Heisenberg Lie algebras as well as sl_2's to generate the algebra. The main difference between Kac-Moody algebras and generalized Kac-Moody algebras is that the roots α of a Kac-Moody algebra may be either real ($(\alpha,\alpha)>0$) or imaginary ($(\alpha,\alpha)\leq 0$) but all the simple roots must be real, while generalized Kac-Moody algebras may also have imaginary simple roots. Kac-Moody algebras have a "denominator formula", which says that a product over positive roots is equal to a sum over the Weyl group; for example, the denominator formula for the affine Kac-Moody algebra $sl_2(\mathbf{R}[z,z^{-1}])$ is the Jacobi triple product identity. Generalized Kac-Moody algebras have a denominator formula which is similar to the one for Kac-Moody algebras, except that it has some extra correction terms for the imaginary simple roots. (The simple roots of a generalized Kac-Moody algebra correspond to a minimal set of generators for the subalgebra corresponding to the positive roots. For Kac-Moody algebras the simple roots also correspond to the points of the Dynkin diagram and to the generators of the Weyl group.)

We return to the monster Lie algebra. This is a generalized Kac-Moody algebra, and we will now write down its denominator formula, which says that a product over the positive roots is a sum over the Weyl group. The positive roots are the vectors (m,n) with $m>0$, $n>0$, and the vector $(1,-1)$, and the root (m,n) has multiplicity $c(mn)$. The Weyl group has order 2 and its nontrivial element maps (m,n) to (n,m), so it exchanges

$p = e^{(1,0)}$ and $q = e^{(0,1)}$. The denominator formula for the monster Lie algebra is the product formula for the j function

$$p^{-1} \prod_{m>0, n\in Z} (1 - p^m q^n)^{c(mn)} = j(p) - j(q).$$

(The left side is antisymmetric in p and q because of the factor of $p^{-1}(1 - p^1 q^{-1})$ in the product.) The reason why we get $j(p)$ and $j(q)$ rather than monomials in p and q on the right hand side (as we would for ordinary Kac-Moody algebras) is because of the correction due to the imaginary simple roots of M. The simple roots of M correspond to a set of generators of the subalgebra E of the elements of M whose degree is to the right of the y axis (so the roots of E are the positive roots of M), and turn out to be the vectors $(1, n)$ each with multiplicity $c(n)$. In the picture of the monster Lie algebra given earlier, the simple roots are given by the column just to the right of the one containing $\mathbf{R}^2$. The sum of the simple root spaces is isomorphic to the space V. The simple root $(1, -1)$ is real of norm 2, and the simple roots $(1, n)$ for $n > 0$ are imaginary of norm $-2n$ and have multiplicity $c(n) = \dim(V_n)$. As these multiplicities are exactly the coefficients of the j function, it is not surprising that j appears in the correction caused by the imaginary simple roots. This discussion is slightly misleading because we have implied that we obtain the product formula of the j function as the denominator formula of the monster Lie algebra by using our knowledge of the simple roots; in fact we really have to use this argument in reverse, using the product formula for the j function in order to work out what the simple roots of the monster Lie algebra are. We do this in section 7.

We can now extract information about the coefficients of the Thompson series $T_g(\tau)$ from a twisted denominator formula for the monster Lie algebra as follows: for an arbitrary generalized Kac-Moody algebra there is a more general version of the Weyl denominator formula which states that

$$\Lambda(E) = H(E),$$

where E is the subalgebra corresponding to the positive roots. Here $\Lambda(E) = \Lambda^0(E) \ominus \Lambda^1(E) \oplus \Lambda^2(E) \ldots$ is a virtual vector space which is the alternating sum of the exterior powers of E, and similarly $H(E)$ is the alternating sum of the homology groups $H_i(E)$ of the Lie algebra E (see [10]). This identity is true for any finite dimensional Lie algebra E because the $H_i(E)$'s are the homology groups of a complex whose terms are the $\Lambda^i(E)$'s. The left hand side corresponds to a product over the positive roots because $\Lambda(A \oplus B) = \Lambda(A) \otimes \Lambda(B)$, $\Lambda(A) = 1 - A$ if A is one dimensional, and E is the sum of $\mathrm{mult}(\alpha)$ one dimensional spaces for each positive root α. For infinite dimensional Lie algebras E we need to be careful that the infinite dimensional virtual vector spaces $H(E)$ and $\Lambda(E)$ are well defined; in this paper they are always differences of graded vector spaces with finite dimensional homogeneous pieces and so are well defined. It is more difficult to identify $H(E)$ with a sum over the Weyl group, and we do this roughly as follows. For Kac-Moody algebras $H_i(E)$ turns out to have dimension equal to the number of elements in the Weyl group of length i; for finite dimensional Lie algebras this was first observed by Bott, and was used by Kostant [26] to give a homological proof of the Weyl character formula. The sum over the homology groups can therefore be identified with a sum over the Weyl group.

For Kac-Moody algebras the same is true and was proved by Garland and Lepowsky [20]. For generalized Kac-Moody algebras things are a bit more complicated. The sum over the homology groups can still be identified with a sum over the Weyl group, but the things we sum are more complicated and contain terms corresponding to the imaginary simple roots.

We can work out the homology groups of E explicitly provided we know the simple roots of our Lie algebra M; for example, the first homology group $H_1(E)$ is the sum of the simple root spaces. For the monster Lie algebra we have worked out the simple roots using its denominator formula, which is the product formula for the j function. In section 8 we use this to work out the homology groups of E, and they turn out to be $H_0(E) = \mathbf{R}$, $H_1(E) = \sum_{n\in\mathbf{Z}} V_n pq^n$, $H_2(E) = \sum_{m>0} V_m p^{m+1}$, and all the higher homology groups are 0. Each homology group is a $\mathbf{Z}^2$-graded representation of the monster, and we use the p's and q's to keep track of the grading. If we substitute these values into the formula $\Lambda(E) = H(E)$ we find that

$$\Lambda(\sum_{n\in\mathbf{Z},m>0} V_{mn}p^m q^n) = \sum_m V_m p^{m+1} - \sum_n V_n pq^n.$$

Both sides of this are virtual graded representations of the monster. If we replace everything by its dimension we recover the product formula for the j function. More generally, we can take the trace of some element of the monster on both sides, which after some calculation gives the identity

$$p^{-1}\exp\Big(-\sum_{i>0}\sum_{m>0,n\in\mathbf{Z}} \mathrm{Tr}(g^i|V_{mn})p^{mi}q^{ni}/i\Big) = \sum_{m\in\mathbf{Z}} \mathrm{Tr}(g|V_m)p^m - \sum_{n\in\mathbf{Z}} \mathrm{Tr}(g|V_n)q^n$$

where $\mathrm{Tr}(g|V_n)$ is the trace of g on the vector space V_n.

These relations between the coefficients $\mathrm{Tr}(g|V_n)$ of the Thompson series are strong enough to determine them from their first few coefficients. Norton and Koike checked that certain modular functions of genus 0 also satisfy the same recursion relations, so we can prove that the Thompson series $T_g(q)$ are these modular functions of genus 0 by checking that the first few coefficients of both functions are the same. Unfortunately this final step of the proof (in section 9) is a case by case check that the first few coefficients are the same. Norton has conjectured [28] that Hauptmoduls with integer coefficients are essentially the same as functions satisfying relations similar to the ones above, and a conceptual proof or explanation of this would be a big improvement to the final step of the proof. (It should be possible to prove Norton's conjecture by a very long and tedious case by case check, because all functions which are either Hauptmoduls or which satisfy the relations above can be listed explicitly. In [1] the authors use a computer to find all "completely replicable" functions with integer coefficients, and they all appear to be Hauptmoduls. Roughly half of them correspond to conjugacy classes of the monster.)

I thank J. McKay, U. Tillman, J. Lepowsky, and the referee, each of whom sent me many useful remarks about a draft of this paper.

Notation. (In roughly alphabetical order.)

$\mathbf{C}$	The complex numbers.
$c(n)$	are the coefficients of the elliptic modular function $j(q)-744$ (defined below).
$c_g(n)$	$=\mathrm{Tr}(g\mid V_n)$ is the n'th coefficient of the Thompson series $T_g(q)$ of g.
$\Gamma_0(N)$	is the subgroup of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $SL_2(\mathbf{Z})$ with $N\mid c$, and $\Gamma_0(N)+$ is its normalizer in $SL_2(\mathbf{R})$; see [13].
δ_i^j	is the Kronecker delta function, which is 1 if $i=j$ and 0 otherwise.
$\Delta(q)$	is the Dedekind delta function $q\prod_{n>0}(1-q^n)^{24}=\eta(q)^{24}$
E	The subalgebra of a generalized Kac-Moody algebra $G=F\oplus H\oplus E$ spanned by the positive root spaces.
E_8	The unique 8-dimensional positive definite even unimodular lattice.
$\epsilon(\alpha)$	The coefficient associated with the root α in the denominator formula of s generalized Kac-Moody algebra. See section 4.
$\eta(q)$	is the Dedekind eta function $q^{1/24}\prod_{n>0}(1-q^n)$. For η_g see sections 9 or 13, and for η_+, η_- see section 11.
$\theta_\Lambda(q)$	$=\sum_{\lambda\in\Lambda}q^{\lambda^2/2}=1+196560q^2+\dots$ is the theta function of the Leech lattice Λ. For other theta functions see section 11.
G	A generalized Kac-Moody algebra; see section 4.
H	The Cartan subalgebra of a generalized Kac-Moody algebra $G=F\oplus H\oplus E$.
$H_i(E)$	is a homology group of the Lie algebra E; $H(E)$ is the alternating sum of the homology groups of E.
$II_{1,1}$	is the unique even 2-dimensional unimodular Lorentzian lattice, which has inner product matrix $\left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right)$. Its elements are usually represented as pairs $(m,n)\in\mathbf{Z}^2=\mathbf{Z}\oplus\mathbf{Z}$, and this element has norm $-2mn$.
$II_{25,1}$	is the unique even 26-dimensional unimodular Lorentzian lattice, which is isomorphic to $\Lambda\oplus II_{1,1}$.
$j(q)$	is the elliptic modular function with $j(q)-744=q^{-1}+196884q+\dots=\theta_\Lambda(q)/\Delta(q)-24=\sum_n c(n)q^n$
Λ	is the Leech lattice, the unique 24 dimensional even unimodular positive definite lattice with no vectors of norm 2. Its elements will often be denoted by λ. For its double cover $\hat{\Lambda}$ see section 12.
$\Lambda(E)$	is the alternating sum of the exterior powers of the vector space E.
$\Lambda^i(E)$	is the i'th exterior power of the vector space E.
M	$=\oplus_{m,n\in\mathbf{Z}}M_{m,n}$ is the monster Lie algebra with root lattice $II_{1,1}$, constructed in section 6.
M_Λ	is the fake monster Lie algebra, whose root lattice is $II_{25,1}$.
$\mu(d)$	is the Moebius function, equal to $(-1)^{(\text{number of prime factors of } d)}$ if d is square free, and 0 otherwise.
$\mathrm{mult}(r)$	The multiplicity of the root r.
p	A formal variable. It can usually be considered as a complex number with $\lvert p\rvert<1$.
$p_{24}(n)$	is the number of partitions of n into parts of 24 colours, so that $\sum_n p_{24}(1+n)q^n=q^{-1}\prod_{n>0}(1-q^n)^{-24}=\Delta(q)^{-1}=q^{-1}+24+324q+3200q^2+25650q^3+176256q^4+1073720q^5+\dots$. These are the multiplicities of roots of the fake monster Lie algebra M_Λ. For p_g see section 14.

q	A formal variable. It can usually be thought of as a complex number with $\lvert q\rvert < 1$, equal to $e^{2\pi i\tau}$. (I.e., the formal series usually converge for $\lvert q\rvert < 1$.)
r^2	The norm (r, r) of the vector r of some lattice.
$\mathbf{R}$	The real numbers.
ρ	is the Weyl vector of a root lattice, which by definition has the property that $(\rho, r) = -(r, r)/2$ for any simple root r. (The Weyl vector is not necessarily in the root lattice, although it does for most of the Lie algebras in this paper.) This has the opposite sign to the usual convention for the Weyl vector, for reasons explained in section 4. For the root lattice $II_{25,1} = \Lambda \oplus II_{1,1} = \Lambda \oplus \mathbf{Z}^2$ of the fake monster Lie algebra M_Λ, ρ is the vector $(0, 0, 1)$.
$\sigma_i(n)$	$= \sum_{d\mid n} d^i$ is the sum of the i'th powers of the divisors of n.
$T_g(q)$	$= \sum c_g(n)q^n$ is the Thompson series of an element g of the monster, with $c_g(n) = \mathrm{Tr}(g\mid V_n)$ where $V = \oplus_{n\in\mathbf{Z}} V_n$ is the module constructed by Frenkel, Lepowsky and Meurman [16,17]. (The spaces V and V_n are denoted by $V^\natural$ and $V^\natural_{-n}$ in [16].)
$\mathrm{Tr}(g\mid U)$	is the trace of an endomorphism g of a vector space U.
τ	A complex number with $\mathrm{Im}(\tau) > 0$.
V	$= \oplus_{n\in\mathbf{Z}} V_n$ is the monster vertex algebra discussed in section 3.
$V_{II_{1,1}}$	is the vertex algebra of the two dimensional even Lorentzian lattice $II_{1,1}$ (or more precisely the vertex algebra of its double cover).
V_Λ	is the the fake monster vertex algebra, which is the vertex algebra of the Leech lattice Λ (or more precisely of its double cover $\hat\Lambda$). See section 12.
W	is a Weyl group. Typical elements are often denoted by w. See section 4.
$\mathbf{Z}$	The integers.
ψ^i	An Adams operation on virtual group representations, defined by $\mathrm{Tr}(g\mid\psi^i(V)) = \mathrm{Tr}(g^i\mid V)$ for V a virtual representation of a group containing g.
ω	A Cartan involution (section 4) or a conformal vector of a vertex algebra (section 3).

2 Introduction (continued).

We describe the results in the second half of the paper (sections 10 to 14). This section can be omitted by those who are only interested in the proof of theorem 1.1.

We construct several Lie superalgebras which are similar to the monster Lie algebra. One method of constructing these is to consider the "twisted" denominator formulas of the monster Lie algebra as untwisted denominator formulas of some other Lie algebras or superalgebras; these seem to be related to other sporadic simple groups. A second method of constructing some of them is to replace the monster vertex algebra V by the vertex algebra of the Leech lattice V_Λ. From this we get the fake monster Lie algebra [8] (where it is called the monster Lie algebra) and several variations of it.

The Lie superalgebras we construct form two families as follows:

(1) A Lie algebra or superalgebra of rank 2 for many conjugacy classes g of the monster simple group. The monster Lie algebra is the one corresponding to the identity element of the monster. The ones corresponding to other elements of the monster are often related to other sporadic simple groups.

(2) A Lie superalgebra for many of the conjugacy classes of the group $\mathrm{Aut}(\hat\Lambda) = 2^{24}.2.Co_1$, where $\hat\Lambda$ is the standard double cover of the Leech lattice Λ (defined in section 12),

$\mathrm{Aut}(\hat\Lambda)$ is the group of its automorphisms which preserve the inner product on Λ, and $Co_1 = \mathrm{Aut}(\Lambda)/\mathbf{Z}_2$ is one of Conway's sporadic simple groups. (The symbol $A.B$ where A and B are groups stands for some extension of B by A, i.e., a group with a normal subgroup A such that the quotient by A is B; the notation is ambiguous.) For example, we get Lie algebras of ranks 26, 18, and 14 corresponding to certain automorphisms of $\hat\Lambda$ of orders 1, 2, and 3, and a Lie superalgebra of rank 10 corresponding to an automorphism of order 2. The Lie algebra of rank 26 is what we now call the fake monster Lie algebra and is studied in [8] (where it is called the monster Lie algebra, because the genuine monster Lie algebra had not been discovered then).

All of these algebras are generalized Kac-Moody algebras or superalgebras. Their root multiplicities can be described explicitly in terms of the coefficients of a finite number of modular forms of weight at most 0. (These modular forms are holomorphic on the upper half plane, and are meromorphic but not necessarily holomorphic at the cusps.) More precisely, there is a sublattice L of finite index in the root lattice, and for each coset of L in the root lattice there is a modular form, which is holomorphic except at the cusps and of weight $1 - \dim(L)/2$, such that the root multiplicity of a vector r in a given coset is the coefficient of $q^{-(r,r)/2}$ of the corresponding modular form. All the algebras have Weyl vectors ρ with norm $\rho^2 = (\rho, \rho) = 0$, and the simple roots are the roots r with $(r, \rho) = -(r, r)/2$. In particular the simple roots and root multiplicities can be described explicitly, and we use this to obtain some infinite product identities from the denominator formulas of these algebras. These algebras are closely related to the sporadic simple groups.

We now discuss both of these families of Lie algebras in more detail. We recall from section 1 that the monster Lie algebra M, is a generalized Kac-Moody algebra whose denominator formula is

$$p^{-1} \prod_{\substack{m>0\\ n\in\mathbf{Z}}} (1 - p^m q^n)^{c(mn)} = j(p) - j(q). \tag{2.1}$$

In section 10 we construct a similar Lie superalgebra for many elements g of the monster. In this case the denominator formula is

$$p^{-1} \prod_{\substack{m>0\\ n\in\mathbf{Z}}} (1 - p^m q^n)^{\mathrm{mult}(m,n)} = T_g(p) - T_g(q) \tag{2.2}$$

where $T_g(q) = \sum \mathrm{Tr}(g|V_n) q^n$ is the Thompson series of the element g, and the multiplicity $\mathrm{mult}(m, n)$ of the root (m, n) is given by

$$\mathrm{mult}(m, n) = \sum_{ds|(m,n)} \mu(d) \mathrm{Tr}(g^s | V_{mn/d^2 s^2})/ds \tag{2.3}$$

where $\mu(d)$ is the Moebius function, equal to $(-1)^{(\text{number of prime factors of } d)}$ if d is square free, and 0 otherwise. (The symbol (m, n) under the summation sign means the greatest common divisor of m and n, rather than the ordered pair.) For example, if N

is a squarefree integer such that the full normalizer $\Gamma_0(N)+$ of $\Gamma_0(N)$ has genus 0 then a special case of (2.2) is

$$p^{-1}\prod_{\substack{m>0\\ n\in\mathbf{Z}}}\prod_{d|(m,n,N)}(1-p^mq^n)^{c_N(mn/d)}=T_N(p)-T_N(q) \tag{2.4}$$

where $T_N(q)=\sum_n c_N(n)q^n$ is the normalized generator of the function field of $\Gamma_0(N)+$ with leading terms $q^{-1}+0+c_N(1)q+\ldots$.

In section 11 to 14 we construct a second series of Lie superalgebras, which are similar to the algebras above except that they are related to the vertex algebra V_Λ of the Leech lattice Λ instead of the monster vertex algebra V. The largest of these, which plays the same role for this series as the monster Lie algebra plays for the previous series, is the algebra which used to be called the monster Lie algebra in [8] and is now called the fake monster Lie algebra. (The Kac-Moody algebra whose Dynkin diagram is that of the reflection group of $II_{25,1}$ has also been called the monster Lie algebra ([7]); it is a large subalgebra of the fake monster Lie algebra, and does not seem to be interesting, except as an approximation to the fake monster Lie algebra.) The root lattice of the fake monster Lie algebra is the 26 dimensional even unimodular Lorentzian lattice $II_{25,1}$, and its denominator formula is

$$e^\rho\prod_{r\in\Pi^+}(1-e^r)^{p_{24}(1-r^2/2)}=\sum_{w\in W}\det(w)w(e^\rho\prod_{n>0}(1-e^{n\rho})^{24}).$$

Here ρ is a norm 0 Weyl vector for the reflection group W of $II_{25,1}$, Π^+ is the set of positive roots, which is the set of vectors r of norm at most 2 which are either positive multiples of ρ or have negative inner product with ρ, $p_{24}(1-r^2/2)$ is the multiplicity of the root r and is equal to the number of partitions of the integer $1-(r,r)/2$ into parts of 24 colours, and the simple roots are the norm 2 vectors r with $(r,\rho)=-1$ together with 24 copies of each positive multiple of ρ. This Lie algebra was first constructed in [3], and the properties stated above were proved in [8]; this construction depended heavily on the ideas in Frenkel [15].

The fake monster Lie algebra is acted on by the group $2^{24}.2.Co_1$ in the same way that the monster Lie algebra is acted on by the monster group, and we construct a superalgebra for many elements of $2^{24}.2.Co_1$ from the fake monster Lie algebra in the same way that we construct a superalgebra for every element of the monster from the monster Lie algebra. Some of the more interesting algebras we get in this way are a fake baby monster algebra of rank 18, a fake Fi_{24} Lie algebra of rank 14 associated with Fischer's sporadic simple group Fi_{24}, a fake Co_1 superalgebra of rank 10 associated with Conway's sporadic simple group Co_1, and several Lie algebras of smaller rank corresponding to some of the other sporadic simple groups involved in the monster.

The superalgebra of rank 10 is particularly interesting, so we describe it explicitly. It is the superalgebra associated with an element g of $\mathrm{Aut}(\hat\Lambda)$ which has order 2 and fixes an 8-dimensional sublattice of Λ, isomorphic to the lattice E_8 with all norms doubled. Its root lattice L is the dual of the sublattice of even vectors of $I_{9,1}$, so that L is a nonintegral lattice of determinant 1/4 all of whose vectors have integral norm. We represent vectors of L as triples (v, m, n), where $v \in E_8$ and $m, n \in \mathbf{Z}$, and (v, m, n) has norm $v^2 - 2mn$. The lattice $I_{9,1}$ is then the set of vectors (v, m, n) with $m + n$ even. We let ρ be the norm 0 vector $(0, 0, 1)$. We let the Weyl group W be the subgroup of $\mathrm{Aut}(L)$ generated by the reflections of norm 1 vectors, so that the simple roots of W are the norm 1 vectors r with $(r, \rho) = -1/2$. (Note that W is not the full reflection group of L, as L also has roots of norm 2.)

The simple roots of our superalgebra are the simple roots of W, together with the positive multiples $n\rho$ $(n > 0)$ of ρ, each with multiplicity $8(-1)^n$. Here we use the convention that multiplicity $-k < 0$ means a superroot of multiplicity k, so that the odd multiples of ρ are superroots. In general the vector (v, m, n) is an ordinary root or a superroot depending on whether $m + n$ is even or odd, so the ordinary roots are those in the sublattice $I_{9,1}$ of L.

The multiplicity of the root $r = (v, m, n) \in L$ is equal to

$$\mathrm{mult}(r) = (-1)^{(m-1)(n-1)} p_g((1 - r^2)/2) = (-1)^{m+n} |p_g((1 - r^2)/2)|,$$

where $p_g(n)$ is defined by

$$\sum p_g(n) q^n = q^{-1/2} \prod_{n>0} (1 - q^{n/2})^{-(-1)^n 8}$$

and as before negative multiplicity means a superroot. The denominator formula for this Lie superalgebra is

$$e^\rho \prod_{r \in \Pi^+} (1 - e^r)^{\mathrm{mult}(r)} = \sum_{w \in W} \det(w) w(e^\rho \prod_{n>0} (1 - e^{n\rho})^{(-1)^n 8}).$$

Similarly the denominator formula for the fake baby monster Lie algebra of rank 18 is

$$e^\rho \prod_{r \in L^+} (1 - e^r)^{p_g(1 - r^2/2)} \prod_{r \in 2L'^+} (1 - e^r)^{p_g(1 - r^2/4)} = \sum_{w \in W} \det(w) w(e^\rho \prod_{i>0} (1 - e^{i\rho})^8 (1 - e^{2i\rho})^8)$$

where L is the Lorentzian lattice which is the sum of the 16-dimensional Barnes-Wall lattice Λ_2 [12] and the two dimensional even Lorentzian lattice $II_{1,1}$, W is its reflection group which has Weyl vector ρ, L' is the dual of L and p_g is defined by $\sum_n p_g(n) q^n = \prod_{n>0} (1 - q^n)^{-8} (1 - q^{2n})^{-8}$. This Lie algebra is associated with an element g of order 2 of $\mathrm{Aut}(\hat{\Lambda})$ which fixes a 16-dimensional lattice of Λ isomorphic to the Barnes-Wall lattice.

There are similar Lie superalgebras associated with many other elements of $\mathrm{Aut}(\hat{\Lambda})$, which have denominator formulas similar to those above.

3 Vertex algebras.

We give a brief summary of the facts about vertex algebras that we use and list the properties of the monster vertex algebra that we need. For more information about vertex algebras, see [3], [9], [16], or [18]; the last two references contain proofs of the results quoted here.

A vertex algebra over the real numbers (which is the only case we use in this paper) is a vector space V over the real numbers with an infinite number of bilinear products, written $u_n v$ for $u, v \in V, n \in \mathbf{Z}$, such that

(1) $u_n v = 0$ for n sufficiently large (depending on u and v).

(2)

$$\sum_{i\in\mathbf{Z}} \binom{m}{i} (u_{q+i}v)_{m+n-i}w = \sum_{i\in\mathbf{Z}} (-1)^i \binom{q}{i} (u_{m+q-i}(v_{n+i}w) - (-1)^q v_{n+q-i}(u_{m+i}w))$$

for all u, v, and w in V and all integers m, n and q.

(3) There is an element $1 \in V$ such that $v_n 1 = 0$ if $n \geq 0$ and $v_{-1}1 = v$.

We often think of u_n as a linear map from V to V, taking v to $u_n v$. In [16,18] these operators are combined into the "vertex operator" $V(u, z) = \sum_{n\in\mathbf{Z}} u_n z^{-n-1}$, which is an operator valued formal Laurent series in z.

For example, if V is a commutative algebra over $\mathbf{R}$ with derivation D then we can make it into a vertex algebra by defining $u_n v$ to be $D^{-n-1}(u)v/(-n-1)!$ for $n < 0$ and 0 for $n \geq 0$, and conversely any vertex algebra over $\mathbf{R}$ for which $u_n v = 0$ whenever $n \geq 0$ arises from a commutative algebra in this way. These are the only finite dimensional examples of vertex algebras, and seem to be the only examples which are easy to construct. The vertex algebras we use here do not have this property and are much harder to construct; for example the detailed construction of the monster vertex algebra in [16] takes up most of a rather long book. Fortunately the details of this construction are not important for this paper, so at the end of this section we list the properties of it that we need.

We define the operator D of a vertex algebra by $D(v) = v_{-2}1$. (In [16] and [18] D is denoted by L_{-1}.) The vector space V/DV is a Lie algebra, where the bracket is defined by $[u, v] = u_0 v$ and DV is the image of V under D.

There is a vertex algebra V_L associated with any even lattice L (or more precisely with a certain central extension $\hat{L}$ of L by a group of order 2), which is constructed in [3]. The underlying vector space of this vertex algebra is the tensor product of the twisted group ring $Q(\hat{L})$ of the double cover $\hat{L}$ of L (see section 12) and the ring of polynomials $S(\oplus_{i>0} L_i)$ over the sum of an infinite number of copies L_i of $L \otimes \mathbf{R}$. We often use the vertex algebra $V_{II_{1,1}}$ of the 2-dimensional even unimodular Lorentzian lattice $II_{1,1}$. There is also a monster vertex algebra V ([3], [16]) which is acted on naturally by the monster sporadic simple group. (The referee has asked me to explain why I have never published my (long and messy) proof of the assertion in [3] that the module V constructed in [17] has the structure of a vertex algebra. The reason is that my proof used many results from the announcement [17], and the only published proof of this announcement, given in the book [16], incorporates the fact that V is a vertex algebra.) The underlying vector space of the monster vertex algebra is a graded vector space $V = \oplus_{n\in\mathbf{Z}} V_n$, with homogeneous pieces of dimensions 1, 0, 196884, ... equal to the coefficients of the elliptic modular function $j(q) - 744$.

A conformal vector of dimension (or "central charge") $c \in \mathbf{R}$ of a vertex algebra V is defined to be an element ω of V such that $\omega_0 v = D(v)$ for any $v \in V$, $\omega_1 \omega = 2\omega$,

$\omega_3\omega = c/2$, $\omega_i\omega = 0$ if $i = 2$ or $i > 3$, and any element of V is a sum of eigenvectors of the operator $L_0 = \omega_1$ with integral eigenvalues. If v is an eigenvector of L_0, then its eigenvalue is called the (conformal) weight of v. If v is an element of the monster vertex algebra V of conformal weight n, we sometimes say that v has degree $n - 1$ $(= n - c/24)$.

The vertex algebra of any c-dimensional even lattice has a canonical conformal vector of dimension c, and the monster vertex algebra has a conformal vector of dimension 24. If ω is a conformal vector of a vertex algebra V then we define operators L_i on V for $i \in \mathbf{Z}$ by

$$L_i = \omega_{i+1}.$$

These operators satisfy the relations

$$[L_i, L_j] = (i-j)L_{i+j} + \binom{i+1}{3}\frac{c}{2}\delta^i_{-j}$$

and so make V into a module over the Virasoro algebra. The operator L_{-1} is equal to D. We define the space P^i to be the space of vectors $w \in V$ such that $L_0(w) = iw$, $L_i(w) = 0$ if $i > 0$. The space $P^1/(DV \cap P^1)$ is a subalgebra of the Lie algebra V/DV, which is equal to P^1/DP^0 for the vertex algebras we use in this paper.

The vertex algebra of any even lattice and the monster vertex algebra both have a real valued symmetric bilinear form (,) such that the adjoint of the operator u_n is $(-1)^i \sum_{j\geq 0} L_1^j(\omega(u))_{2i-j-n-2}/j!$ if u has degree i, where ω is the automorphism of the vertex algebra defined by $\omega(e^w) = (-1)^{(w,w)/2}(e^w)^{-1}$ for e^w an element of the twisted group ring of L corresponding to the vector $w \in L$, or is 1 in the case of the monster vertex algebra. (There is an unfortunate clash of notation here, because ω is the standard notation for both the conformal vector and the Cartan automorphism of a Lie algebra. If L is the root lattice of a finite dimensional Lie algebra of type A_n, D_n, E_6, E_7, or E_8 then the Lie algebra is a subalgebra of the vertex algebra of L and the involution ω defined above restricts to the Cartan automorphism of the Lie algebra.) If a vertex algebra has a bilinear form with the properties above we say that the bilinear form is compatible with the conformal vector.

Frenkel, Lepowsky, and Meurman [16] use "vertex operator algebras", rather than vertex algebras, so we explain what the difference is. A vertex operator algebra is a vertex algebra with a conformal vector such that the eigenspaces of L_0 are all finite dimensional with nonnegative integral eigenvalues. For example, the monster vertex algebra is a vertex operator algebra. (Its conformal vector spans the subspace of V_1 fixed by the monster.) The vertex algebras used in this paper all have conformal vectors but the eigenspaces of L_0 are not always finite dimensional. An example of a vertex algebra without these properties is the vertex algebra $V_{II_{1,1}}$ of the 2-dimensional even unimodular Lorentzian lattice $II_{1,1}$.

If V and W are vertex algebras, then their tensor product $V \otimes W$ as vector spaces is also a vertex algebra, if $(a \otimes b)_n(c \otimes d)$ is defined to be $\sum_{i\in\mathbf{Z}}(a_ic) \otimes (b_{n-1-i}d)$ and the identity 1 is defined to be $1 \otimes 1$. (See [18].) If V and W have conformal vectors ω_V and ω_W of dimensions m and n then $\omega_V \otimes 1 + 1 \otimes \omega_W$ is a conformal vector of $V \otimes W$ of dimension $m + n$. If V and W have bilinear forms compatible with the conformal vectors then so does $V \otimes W$; in fact the obvious bilinear form on the tensor product will do.

The monster vertex algebra V [3, 16] is a vertex algebra acted on by the monster. The construction given in [16] is rather long, but fortunately we do not need to know many details of its construction. We list the three properties of it we do use, in case anyone finds a simpler construction of it that they wish to prove also satisfies Conway and Norton's moonshine conjectures.

(1) V is a vertex algebra over $\mathbf{R}$ with a conformal vector ω of dimension 24 and a positive definite bilinear form such that the adjoint of u_n is given by the expression above (with the automorphism ω acting trivially).

(2) V is a sum of eigenspaces V_i of the operator L_0, where V_i is the eigenspace on which L_0 has eigenvalue $i+1$, and the dimension of V_i is given by $\sum \dim(V_i)q^i = j(q) - 744 = q^{-1} + 196884q + \ldots$.

These two conditions turn out to imply that if g is any automorphism of V preserving the vertex algebra structure, the conformal vector and the bilinear form then $\sum \mathrm{Tr}(g|V_i)q^i$ is a completely replicable function (which means that identity (8.3) holds).

(3) The monster simple group acts on V, preserving the vertex algebra structure, the conformal vector ω and the bilinear form. The first few representations V_i of the monster (after $V_{-1} = \chi_1$, $V_0 = 0$) decompose as $V_1 = \chi_1 + \chi_2$, $V_2 = \chi_1 + \chi_2 + \chi_3$, $V_3 = 2\chi_1 + 2\chi_2 + \chi_3 + \chi_4$, $V_5 = 4\chi_1 + 5\chi_2 + 3\chi_3 + 2\chi_4 + \chi_5 + \chi_6 + \chi_7$ where χ_i, $1 \leq i \leq 7$ are the first seven irreducible representations of the monster, indexed in order of increasing dimension.

The construction of the vector space V and the action of the monster on it, preserving a small part of the vertex algebra structure, were announced in [17], and the vertex algebra structure was announced in [3]. The results quoted above have all been proved in [16] and [18], apart from the explicit description of the representations in (3). We show in section 9 that condition (3) can be proved using the explicit description in [16] of the traces of certain elements of the monster acting on V. In the course of proving theorem 1.1 we show that these three conditions characterize V as a graded representation of the monster.

4 Generalized Kac-Moody algebras

We summarize the results about generalized Kac-Moody algebras that we use, which can be found in [4], [5], [20], and the third edition of Kac's book [23]. We modify the original definition of generalized Kac-Moody algebras in [4] slightly so that these algebras are closed under taking universal central extensions, as in [5]. (This is not necessary for the proof of theorem 1.1, but makes the converse to theorem 4.1 slightly neater.) All Lie algebras are Lie algebras over the reals.

A Lie algebra G is defined to be a (split) generalized Kac-Moody algebra if it has an almost positive definite contravariant bilinear form, which means that G has the following three properties.

1. G can be $\mathbf{Z}$-graded as $G = \oplus_{i\in\mathbf{Z}} G_i$ and G_i is finite dimensional if $i \neq 0$.
2. G has an involution ω which maps G_i into G_{-i} and acts as -1 on G_0, so in particular G_0 is abelian.
3. G has an invariant bilinear form (,) invariant under ω such that G_i and G_j are orthogonal if $i \neq -j$, and such that $-(g, \omega(g)) > 0$ if g is a nonzero homogeneous element of G of nonzero degree.

The bilinear form $(,)_0$ defined by $(a,b)_0 = -(a,\omega(b))$ is called the contravariant bilinear form of G, and is positive definite on G_i if $i \neq 0$. For example any finite dimensional split semisimple Lie algebra over the reals has these properties, with ω equal to a Cartan involution, (,) equal to the Killing form, and the grading determined by the eigenvalues of some suitably normalized regular element. (If the contravariant bilinear form on G is positive semidefinite, then G is essentially a sum of affine, Heisenberg, and finite dimensional simple split Lie algebras.) In general the contravariant form can be indefinite on G_0. For the algebras we use in this paper the subalgebra G_0 has a vector such that the contravariant form is positive semidefinite on its orthogonal complement.

Suppose that $a_{ij}, i,j \in I$ is a symmetric countable (possibly infinite) real matrix such that $a_{ij} \leq 0$ if $i \neq j$ and such that if $a_{ii} > 0$ then $2a_{ij}/a_{ii}$ is an integer for any j. Then the universal generalized Kac-Moody algebra G of this matrix is defined to be the Lie algebra generated by elements e_i, f_i, h_{ij} for $i,j \in I$ satisfying the following relations. (It is true but not obvious that any universal generalized Kac-Moody algebra is a generalized Kac-Moody algebra; this fact is not needed in this paper.)

1. $[e_i, f_j] = h_{ij}$
2. $[h_{ij}, e_k] = \delta_i^j a_{ik} e_k, [h_{ij}, f_k] = -\delta_i^j a_{ik} f_k$
3. If $a_{ii} > 0$ and $i \neq j$ then $\mathrm{ad}(e_i)^n e_j = \mathrm{ad}(f_i)^n f_j = 0$, where $n = 1 - 2a_{ij}/a_{ii}$.
4. If $a_{ii} \leq 0, a_{jj} \leq 0$ and $a_{ij} = 0$ then $[e_i, e_j] = [f_i, f_j] = 0$.

We often write h_i for h_{ii}. There is a unique invariant bilinear form on this algebra such that $(e_i, f_j) = \delta_i^j$; this implies that $(h_i, h_j) = a_{ij}$.

Remark. If $a_{ii} > 0$ for all $i \in I$ then this algebra is the same as the Kac-Moody algebra with symmetrized Cartan matrix a_{ij}. In general these algebras have almost all the properties that Kac-Moody algebras have, and the only major difference is that generalized Kac-Moody algebras are allowed to have imaginary simple roots.

The main theorem about generalized Kac-Moody algebras [4,5] says that they can all be obtained from universal generalized Kac-Moody algebras.

Theorem 4.1. *Suppose that G is a generalized Kac-Moody algebra with some given grading, involution ω and bilinear form (,). Then there is a unique universal generalized Kac-Moody algebra, graded by putting* $\deg(e_i) = -\deg(f_i) = n_i$ *for some positive integers n_i, with a homomorphism f (not necessarily unique) to G such that*

(1) f preserves the gradings, involutions and bilinear forms (where the universal generalized Kac-Moody algebra is given the grading, involution and bilinear form defined above).

(2) The kernel of f is in the centre of the universal generalized Kac-Moody algebra (which is contained in the abelian subalgebra spanned by the elements h_{ij}).

(3) The image of f is an ideal of G, and G is the semidirect product of this subalgebra and a subalgebra of the abelian subalgebra G_0. Moreover the images of all the generators e_i and f_i are eigenvectors of G_0.

Remark. The converse of this theorem is also true (but not needed here): any universal generalized Kac-Moody algebra with a countable number of simple roots is a generalized Kac-Moody algebra. (The proof of this converse in [4] contains a gap which was pointed out and corrected by Kac in [23].) Theorem 4.1 says that we can construct any generalized

Kac-Moody algebra from some universal generalized Kac-Moody algebra by factoring out some of the centre and adding a commuting algebra of outer derivations.

We list some of the properties of universal generalized Kac-Moody algebras G from [4] and [5], many of which are also proved in the third edition of [23].

(1) The element h_{ij} is 0 unless the i'th and j'th columns of a are equal. The elements h_{ij} for which the i'th and j'th columns of a are equal form a basis for an abelian subalgebra H of G, called its Cartan subalgebra. In the case of Kac-Moody algebras, the i'th and j'th columns of a cannot be equal unless $i = j$, so the only nonzero elements h_{ij} are those of the form h_{ii}, which are usually denoted by h_i. Every nonzero ideal of G has nonzero intersection with H. The centre of G is contained in H and contains all the elements h_{ij} for $i \neq j$.

(2) The root lattice is defined to be the free abelian group generated by elements r_i for $i \in I$, with the bilinear form given by $(r_i, r_j) = a_{ij}$. The elements r_i are called the simple roots. The universal generalized Kac-Moody algebra is graded by the root lattice if we let e_i have degree r_i and f_i have degree $-r_i$. There is a natural homomorphism of abelian groups from the root lattice to the Cartan subalgebra H taking r_i to h_i which preserves the bilinear forms. If r is in the root lattice then the vector space of elements of the Lie algebra of that degree is called the root space of r; if r is nonzero and has a nonzero root space then r is called a root of the generalized Kac-Moody algebra. A root r is called positive if it is a sum of simple roots, and negative otherwise. It is called real if it has positive norm (r, r) and imaginary otherwise.

(3) The Weyl vector ρ is the additive map from the root lattice to $\mathbf{R}$ taking r_i to $-(r_i, r_i)/2$ for all $i \in I$. (This is the opposite sign from the usual definition.) We write its value on a root r as (ρ, r). The fundamental Weyl chamber (which we call the Weyl chamber for short, because it is the only one we use) is the set of vectors v of the Cartan subalgebra H with $(v, r_i) \leq 0$ for all real simple roots r_i, where we identify real simple roots with their images in H. (There is a misprint in the definition of Weyl chamber in [4]; the word "real" in "real simple roots" was omitted.)

Remark. The reason why we use nonstandard sign conventions for the Weyl vector and Weyl chamber is that we are using the convention that "everything should be in the Weyl chamber if possible". We want the imaginary simple roots to be in the Weyl chamber which forces us to use the opposite Weyl chamber from usual; similarly we change the sign of the Weyl vector to get it (or rather its image in H, when this exists) in the Weyl chamber, and we use lowest weight modules (rather than highest weight modules) because their lowest weights are in the Weyl chamber. For most of the generalized Kac-Moody algebras in this paper the Cartan subalgebra has a Lorentzian bilinear form, so that the vectors of norm at most 0 form two solid cones; our sign conventions usually make all the interesting vectors of norm at most 0 lie in the same solid cone.

(4) If a has no zero columns then G is perfect and equal to its own universal central extension. If a is not the direct sum of two smaller matrices, not the 1×1 zero matrix, and not the matrix of an affine Kac-Moody algebra, then G modulo its centre is simple.

(5) Suppose we choose a positive integer n_i for each $i \in I$. Then we can grade G by putting $\deg(e_i) = -\deg(f_i) = n_i$. The degree 0 piece of G is the Cartan subalgebra H.

(6) G has an involution ω with $\omega(e_i) = -f_i, \omega(f_i) = -e_i$, called the Cartan involution.

There is a unique invariant bilinear form $(,)$ on G such that $(e_i, f_i) = 1$ for all i, and it also has the property that $-(g, \omega(g)) > 0$ whenever g is a homogeneous element of nonzero degree. In particular the universal generalized Kac-Moody algebra of any matrix as above is a generalized Kac-Moody algebra.

(7) There is a character formula for the simple lowest weight module M_λ of G with a lowest weight λ in the real vector space of the root lattice such that $(\lambda, r_i) \leq 0$ for $i \in I$ and $2(\lambda, r_i)/(r_i, r_i)$ is integral if the simple root r_i is real. This states that

$$Ch(M_\lambda)e^\rho \prod_{\alpha>0}(1-e^\alpha)^{\mathrm{mult}(\alpha)} = \sum_{w\in W} \det(w) w(e^\rho \sum_\alpha \epsilon(\alpha) e^{\alpha+\lambda})$$

The only case of this we need is for the trivial one dimensional module with $\lambda = 0$, when it becomes the denominator formula

$$e^\rho \prod_{\alpha>0}(1-e^\alpha)^{\mathrm{mult}(\alpha)} = \sum_{w\in W} \det(w) w(e^\rho \sum_\alpha \epsilon(\alpha) e^{\alpha})$$

Here ρ is the Weyl vector, $\alpha > 0$ means α is a positive root, and W is the Weyl group, which is the group of isometries of the root lattice generated by the reflections corresponding to the real simple roots. If $w \in W$ then $\det(w)$ is defined to be $+1$ or -1 depending on whether w is the product of an even or odd number of reflections; if the root lattice is finite dimensional this is just the usual determinant of w acting on it. We define $\epsilon(\alpha)$ for α in the root lattice to be $(-1)^n$ if α is the sum of a set of n pairwise orthogonal imaginary simple roots that are all orthogonal to λ, and 0 otherwise. If G is a Kac-Moody algebra then there are no imaginary simple roots so the sum over α is 1 and we recover the usual character and denominator formulas.

If we have a generalized Kac-Moody algebra G then we choose some universal generalized Kac-Moody algebra mapping to G as in theorem 4.1, and call the Weyl group, root lattice, and so on of the universal generalized Kac-Moody algebra the Weyl group, root lattice and so on of G.

For the generalized Kac-Moody algebras in this paper, the Cartan subalgebra H is usually finite dimensional and its bilinear form is nonsingular. We often think of roots of G as vectors of H by using the natural map from the root lattice to H, taking r_i to h_i. We will often abuse terminology and call the corresponding vectors of H roots. The map from the set of roots to H is not usually injective, so we need to make a few comments to clarify what happens. It is possible for $n > 1$ imaginary simple roots to have the same image r in H; when this happens we say that r is a simple root of multiplicity n. (In general it is possible for the same vector of H to be the image of simple and nonsimple roots, but this does not happen for the Lie algebras in this paper.) The simple roots in H are usually linearly dependent so in general there is no reason why there should exist a Weyl vector ρ in H such that $(\rho, r_i) = -r_i^2/2$ for all real simple roots r_i; however such a vector ρ does exist for the Lie algebras in this paper.

The denominator formula also needs an obvious modification if we think of the roots as elements of H. If a simple root has multiplicity $n > 1$ then the "set" of simple roots should include n copies of it, i.e., the "set" of simple roots is really a "multiset" rather than a set. (A multiset is a "set which may contain several copies of the same object", or more precisely a map from a set to the positive integers.) The expression $\epsilon(\alpha)$ in the character formula is then the sum of terms $(-1)^n$ for all ways of writing α as a sum of a "set" (or more precisely a multiset) of pairwise orthogonal imaginary simple roots that are all orthogonal to λ. For example, if there is only one simple root r of norm 0 and multiplicity n, then $\epsilon(ir) = \binom{n}{i}(-1)^i$ because there are $\binom{n}{i}$ ways of writing ir as a sum of a multiset of pairwise orthogonal simple roots.

Remark. Most of these results can be extended to superalgebras with little difficulty, provided that we interpret roots of negative multiplicity as being "superroots". The only extra detail worth pointing out is that we have to add the condition that all simple superroots are imaginary (and not real). For example, many of the finite dimensional superalgebras are generalized Kac-Moody superalgebras with simple superroots of norm 0. Warning: The Cartan matrix of a finite dimensional superalgebra may depend on the $\mathbf{Z}$-grading chosen.

Any generalized Kac-Moody algebra can be written as a direct sum $E\oplus H\oplus F$ of subalgebras, where H is the Cartan subalgebra and E and F are the subalgebras corresponding to the positive and negative roots. The standard sequence

$$\ldots \to \Lambda^2(E) \to \Lambda^1(E) \to \Lambda^0(E) \to 0$$

whose homology groups are those of the Lie algebra E [10] shows that $\Lambda(E) = H(E)$, where $\Lambda(E)$ is the virtual vector space $\Lambda^0(E)\ominus\Lambda^1(E)\oplus\Lambda^2(E)\ldots$ formed from the alternating sum of the exterior powers of E, and $H(E)$ is the virtual vector space $H_0(E)\ominus H_1(E)\oplus H_2(E)\ldots$ formed from the alternating sum of the homology groups of E with real coefficients. If L is the root lattice of M, then both sides are L-graded virtual vector spaces whose homogeneous pieces are finite dimensional, so these infinite sums are meaningful.

The denominator formula for any generalized Kac-Moody algebra follows from the fact that $\Lambda(E) = H(E)$ as virtual L-graded vector spaces. The character of $\Lambda(E)$ is a product over positive roots. In the case of ordinary Kac-Moody algebras the vector space $H_i(E)$ has a basis corresponding to the elements of the Weyl group of length i, so the character of $H(E)$ is a sum over the Weyl group. If we work out everything explicitly we obtain the denominator formula

$$e^\rho \prod_{\alpha>0}(1-e^\alpha)^{\mathrm{mult}(\alpha)} = \sum_{w\in W}\det(w)w(e^\rho)$$

for Kac-Moody algebras, and a similar but slightly more complicated calculation gives the denominator formula for generalized Kac-Moody algebras. For Kac-Moody algebras this calculation is carried out in Garland and Lepowsky [20] and most of their methods and results apply to generalized Kac-Moody algebras with only minor changes. They obtain a more general result which not only gives the character formula of any highest weight module, but also works if E is replaced by various other radicals of parabolic subalgebras. We do not need these more general results, but they are easily extended to the case of generalized Kac-Moody algebras.

The groups $H_i(E)$ can be calculated for generalized Kac-Moody algebras in the same way that they are calculated for ordinary Kac-Moody algebras in Garland and Lepowsky

[20]. The result we get is that $H_i(E)$ is the subspace of $\Lambda^i(E)$ spanned by the homogeneous vectors of $\Lambda^i(E)$ whose degrees $r \in L$ satisfy $(r+\rho)^2 = \rho^2$. The degree r of any homogeneous vector of $\Lambda^i(E)$ satisfies $(r+\rho)^2 \leq \rho^2$, so $H_i(E)$ can be thought of as a sort of boundary of $\Lambda^i(E)$. These homology groups are easy to work out for any generalized Kac-Moody algebra, provided that we know the simple roots and their multiplicities. For the monster Lie algebra, we carry out this calculation in section 8.

In practice it is sufficient to know the subspace of $H(E)$ of elements whose degree r has the property that $r+\rho$ is in the fundamental Weyl chamber. This subspace is isomorphic to the subspace of $\Lambda(E)$ of all elements that can be written in the form $e_1 \wedge e_2 \wedge \ldots$ where the e_i's are vectors in the root spaces of pairwise orthogonal imaginary simple roots, so that the character of this subspace is just the sum $\sum_\alpha \epsilon(\alpha)e^\alpha$ that appears in the denominator formula. There is a similar description of the homology groups of E with coefficients in any lowest weight module which can be used to prove the character formula for lowest weight modules, but we do not use this.

5 The no-ghost theorem

We prove a slight extension of the no-ghost theorem. The idea of using the no-ghost theorem to prove results about Kac-Moody algebras appears in Frenkel's paper [15], which also contains a proof of the no-ghost theorem. The original proof of Goddard and Thorn [21] works for the cases we need with only trivial modifications. For convenience we give a quick sketch of their proof.

Theorem 5.1. *(The no-ghost theorem.) Suppose that V is a vector space with a nonsingular bilinear form (,) and suppose that V is acted on by the Virasoro algebra in such a way that the adjoint of L_i is L_{-i}, the central element of the Virasoro algebra acts as multiplication by 24, any vector of V is a sum of eigenvectors of L_0 with nonnegative integral eigenvalues, and all the eigenspaces of L_0 are finite dimensional. We let V^i be the subspace of V on which L_0 has eigenvalue i. Assume that V is acted on by a group G which preserves all this structure. We let $V_{II_{1,1}}$ be the vertex algebra of the double cover $\hat{I}I_{1,1}$ of the two dimensional even unimodular Lorentzian lattice $II_{1,1}$ (so that $V_{II_{1,1}}$ is $II_{1,1}$-graded, has a bilinear form (,), and is acted on by the Virasoro algebra as in section 3). We let P^1 be the subspace of the vertex algebra $V \otimes V_{II_{1,1}}$ of vectors v with $L_0(v) = v, L_i(v) = 0$ for $i > 0$, and we let P^1_r be the subspace of P^1 of degree $r \in II_{1,1}$. All these spaces inherit an action of G from the action of G on V and the trivial action of G on $V_{II_{1,1}}$ and $\mathbf{R}^2$. Then the quotient of P^1_r by the nullspace of its bilinear form is naturally isomorphic, as a G module with an invariant bilinear form, to $V^{1-(r,r)/2}$ if $r \neq 0$ and to $V^1 \oplus \mathbf{R}^2$ if $r = 0$.*

The name "no-ghost theorem" comes from the fact that in the original statement of the theorem in [21], V was part of the underlying vector space of the vertex algebra of a positive definite lattice so the inner product on V^i was positive definite, and thus P^1_r had no ghosts (i.e. vectors of negative norm) for r nonzero. The space V^i is the space V_{i-1} used in the rest of this paper.

Sketch of proof (taken from Goddard and Thorn, [21]). Fix some nonzero $r \in II_{1,1}$ and some norm 0 vector w in $II_{1,1}$ with $(r,w) \neq 0$.

We use the following operators. We have an action of the Virasoro algebra on $V \otimes V_{II_{1,1}}$ generated by its conformal vector. The operators L_i of the Virasoro algebra satisfy the relations

$$[L_i, L_j] = (i-j)L_{i+j} + 26(i^3 - i)\delta^i_{-j}/12.$$

and the adjoint of L_i is L_{-i}. (The 26 comes from the 24 in theorem 5.1 plus the dimension of $II_{1,1}$.) We define operators K_i, $i \in \mathbf{Z}$ by $K_i = v_{i-1}$ where v is the element $e^{-w}_{-2}e^w$ of the vertex algebra of $II_{1,1}$, and e^w is an element of the group ring of the double cover of $II_{1,1}$ corresponding to $w \in II_{1,1}$ and e^{-w} is its inverse. (There are in fact two possible choices for e^w which differ by a factor of -1, but it does not matter which we choose because this factor of -1 cancels out in the expression $e^{-w}_{-2}e^w$.) These operators satisfy the relations

$$\begin{aligned}[L_i, K_j] &= -jK_{i+j}\\ [K_i, K_j] &= 0\end{aligned}$$

because w has norm 0, and the adjoint of K_i is K_{-i}.

We define the following subspaces of $V \otimes V_{II_{1,1}}$.

H is the subspace of $V \otimes V_{II_{1,1}}$ of degree $r \in II_{1,1}$. H^1 is its subspace of vectors h with $L_0(h) = h$.
P, is the subspace of H of all vectors h with $L_i(h) = 0$ for $i > 0$. $P^1 = H^1 \cap P$.
S, the space of spurious vectors, is the subspace of H of vectors perpendicular to P. $S^1 = H^1 \cap S$.
N $= S \cap P$ is the radical of the bilinear form of P, and $N^1 = H^1 \cap N$.
T, the transverse space, is the subspace of P annihilated by all the operators $K_i, i > 0$, and $T^1 = H^1 \cap T$.
K is the space generated by the action of the operators $K_i, i \in \mathbf{Z}$ on T.
Ve^r is the subspace $V \otimes e^r$ of H.

We have the following inclusions of subspaces of H:

$$\begin{tikzcd}
S & & P & & K & \\
& N \arrow[ul] \arrow[ur] & & T \arrow[ul] \arrow[ur] & & Ve^r \arrow[ul]
\end{tikzcd}$$

and we construct the isomorphism from $V^{1-(r,r)/2}$ to $P^1/N \cap P^1$ by zigzagging up and down this diagram; more precisely we show that Ve^r and T are both isomorphic to K modulo its nullspace, and then we show that T^1 is isomorphic to P^1 modulo its nullspace $P^1 \cap N$.

Lemma 5.1. *If f is a vector of nonzero norm in T then the vectors of the form*

$$L_{m_1}L_{m_2}\ldots K_{n_1}K_{n_2}\ldots(f)$$

for all sequences of integers with $0 > m_1 \geq m_2 \ldots$, $0 > n_1 \geq n_2 \ldots$ are linearly independent and span a space invariant under the operators K_i and L_i on which the bilinear form is nonsingular.

Sketch of proof: It is possible to order these elements so that the matrix of inner products is upper triangular with nonzero diagonal elements. This implies that they are independent and span a space on which the inner product is nonsingular. The other statements in the lemma are easy to prove.

Lemma 5.2. *The bilinear form on T is nonsingular, and K is the direct sum of T and the nullspace of K.*

Proof. If we choose an orthogonal basis for T, then the set of all vectors generated from them as in Lemma 5.1 is a basis for H. As H is nonsingular, T must therefore also be nonsingular. It now follows that K is the direct sum of its nullspace and T, because K is generated from T by the operators K_i and anything generated from T by these operators is in the nullspace of K.

Lemma 5.3. *Ve^r is naturally isomorphic to T.*

Proof. K is also the subspace of H annihilated by all the operators K_i, $i > 0$ from which it follows that K is the direct sum of Ve^r and the nullspace of K. By lemma 5.2 K is also the direct sum of T and the nullspace of K, so this gives a natural isomorphism from Ve^r to T, because they are both isomorphic to the quotient of K by its nullspace.

Lemma 5.4. *The associative algebra generated by the elements L_i for $i < 0$ is generated by elements mapping S^1 into S.*

Proof. Calculation shows that the operators L_{-1} and $L_{-2}+3L_{-1}^2/2$ have this property, and they generate the algebra generated by all L_i's with $i < 0$. (It is in doing these calculations that we need to use the fact that the center of the Virasoro algebra acts as 24 on V and therefore as 26 on H.)

Lemma 5.5. *P^1 is the direct sum of T^1 and N^1.*

Proof. We have to show that any vector p of P^1 can be written as $p = t+n$ for $t \in T^1$, $n \in N^1$. By lemma 5.3, $p = k + s$ for some unique $k \in K^1$, $s \in S^1$ and by lemma 5.4 k and s are both mapped into K and S by some set of generators of the algebra generated by the L_i's, $i > 0$. As these generators annihilate p, they must also annihilate k and s. Hence k is in $K^1 \cap P^1 = T^1$ and s is in $S^1 \cap P^1 = N^1$, so that $P^1 = T^1 \oplus N^1$.

The no ghost theorem now follows immediately, because by lemma 5.3 $V^{1-(r,r)/2}e^r$ is isomorphic to T^1, and by lemma 5.5 T^1 is isomorphic to the quotient of P^1 by its nullspace.

6 Construction of the Monster Lie algebra.

We construct the monster Lie algebra M in the proof of the following theorem. This is a generalized Kac-Moody algebra acted on by the monster simple group and graded by $II_{1,1} = \mathbf{Z}^2$. Recall that V is the vertex algebra of the monster, so that V is graded with pieces of dimensions 1, 0, 196884, ... , and $V_{II_{1,1}}$ is the vertex algebra of the two dimensional even Lorentzian lattice $II_{1,1}$ with matrix $\left(\begin{smallmatrix}0 & -1\\ -1 & 0\end{smallmatrix}\right)$.

Theorem 6.1. *There exists a Lie algebra M, which we will call the monster Lie algebra, with the following properties.*

(1) M is $II_{1,1}$-graded.

(2) M has a contravariant bilinear form $(,)_0$ which is positive definite on the piece of degree $(m,n) \neq (0,0)$. (By "contravariant bilinear form" we mean that the $II_{1,1}$-graded Lie algebra M has an involution ω which acts as -1 on $II_{1,1}$ and as -1 on

the piece of degree (0,0), such that the form $(u, v) = -(u, \omega(v))_0$ is invariant, and $(u, v) = 0$ unless $\deg(u) + \deg(v) = 0$.)

(3) M is acted on by the monster. As a representation of the monster, the piece of M of degree (m, n) is isomorphic to V_{mn} if $(m, n) \neq (0, 0)$, and to the trivial representation $\mathbf{R}^2$ if $(m, n) = (0, 0)$ where V_n is the piece of the monster vertex algebra V of conformal weight $n + 1$.

Proof. The tensor product $V \otimes V_{II_{1,1}}$ of the vertex algebras V and $V_{II_{1,1}}$ is also a vertex algebra. If P^i is the space of vectors v of this vertex algebra satisfying $L_j(v) = 0$ if $j > 0$, $L_0(v) = v$, then P^1/DP^0 is a Lie algebra with an invariant bilinear form (,), and an involution ω induced by the trivial automorphism of V and the involution ω of $V_{II_{1,1}}$ defined in section 3. We define the monster Lie algebra M to be the quotient of the Lie algebra P^1/DP^0 by the kernel of the form (,). (The kernel of the bilinear form on P^1 is strictly larger than DP^0.)

The $II_{1,1}$ grading of the vertex algebra $V_{II_{1,1}}$ induces a $II_{1,1}$ grading on the Lie algebra M. The no-ghost theorem 5.1 implies that the piece of M of degree $(m, n) \in II_{1,1}$ is isomorphic to the piece of V of degree $1 - (m, n)^2/2 = 1 - mn$ if $v \neq 0$ and to $\mathbf{R}^2$ if $v = 0$, and that $(g, \omega(g)) > 0$ if $g \in M$ is nonzero and homogeneous of nonzero degree in $II_{1,1}$. This proves theorem 6.1.

Remark. The construction of a Lie algebra from a vertex algebra in theorem 6.1 can be carried out for any vertex algebra V with a conformal vector, but it is only when this vector has dimension 24 that we can apply the no-ghost theorem to identify the homogeneous pieces of M with those of V. The important point is that the bilinear form on M is positive definite on any piece of nonzero degree, and this need not be true if the conformal vector has dimension greater than 24, even if the inner product on V is positive definite.

Theorem 6.2. *The monster Lie algebra M is a generalized Kac-Moody algebra.*

Proof: This follows from theorem 6.1. The only condition in the definition of a generalized Kac-Moody algebra that is not immediately obvious from the conclusion of theorem 6.1 is the one about the $\mathbf{Z}$-grading of M. We can $\mathbf{Z}$-grade M in a suitable way by letting elements of M of degree $(m, n) \in II_{1,1}$ have degree $2m + n \in \mathbf{Z}$.

Remark. If V_L is the vertex algebra of any positive definite even lattice L of dimension at most 24 then the bilinear form on the corresponding Lie algebra M_L is still almost positive definite, so M_L is a generalized Kac-Moody algebra. If V_Λ is the vertex algebra of the Leech lattice Λ then M_Λ is (by definition) the fake monster Lie algebra, and its simple roots are described explicitly in [8] (where it is called the monster Lie algebra, or more precisely proved to be isomorphic to something called the monster Lie algebra there). Section 12 contains a summary of results about the fake monster Lie algebra. I have not found any other lattices for which the simple roots of the corresponding Lie algebra can be found explicitly, although for the lattice $E_8 \oplus E_8$ all the simple roots of norm 0, -2, or -4 have multiplicity 0 or 1 so there may be something interesting going on in this case. There are some calculations for the case when V_L is the vertex algebra of $L = E_8$ in Kac, Moody and Wakimoto [24].

The monster Lie algebra can also be constructed as a semi-infinite cohomology group; see [19].

The construction of the monster Lie algebra above looks bizarre at first sight, so we briefly explain some of the motivation behind it. The fake monster vertex algebra M_Λ [8] is the Lie algebra P^1/(kernel of bilinear form) of the vertex algebra of the lattice $II_{25,1}$. This lattice is the sum of the Leech lattice Λ and the lattice $II_{1,1}$, and the vertex algebra of the sum of two lattices Λ and $II_{1,1}$is the tensor product of the vertex algebras V_Λ and $V_{II_{1,1}}$ of the lattices, so the fake monster Lie algebra is the space P^1/kernel of $V_\Lambda \otimes V_{II_{1,1}}$. The monster vertex algebra V is similar to the vertex algebra V_Λ, which suggests replacing V_Λ by V in the construction above, and this gives the monster Lie algebra M.

7 The simple roots of the monster Lie algebra.

We find the simple roots of the monster Lie algebra M constructed in section 6. This algebra turns out to have a Weyl vector, i.e., a vector ρ with $(\rho, r) = -(r,r)/2$ for all simple roots r.

We find the simple roots of the monster Lie algebra by using the following identity, which turns out to be its denominator formula.

Lemma 7.1.

$$p^{-1}\prod_{\substack{m>0\\ n\in\mathbf{Z}}}(1-p^mq^n)^{c(mn)} = j(p)-j(q) \tag{7.1}$$

where $j(q)-744 = \sum_n c(n)q^n = q^{-1}+196884q+\ldots$.

Proof. This is essentially just lemma 2 of section 4 of [8]. For convenience we recall the proof. If we multiply the left hand side by p and take its logarithm we get

$$\begin{aligned}
&-\sum_{m>0}\sum_{n\in\mathbf{Z}}\sum_{k>0} c(mn)p^{mk}q^{nk}/k\\
=&-\sum_{m>0}\sum_{n\in\mathbf{Z}}\sum_{0<k|(m,n)}\frac{1}{k}c(\frac{mn}{k^2})p^mq^n\\
=&-\sum_{m>0}T_m(\sum_{n\in\mathbf{Z}}c(n)q^n)p^m\\
=&-\sum_{m>0}T_m(j(q)-744)p^m\\
=&\sum_{m>0}f_m(q)p^m
\end{aligned}$$

where T_m is the m'th Hecke operator (see Serre [30] proposition 12, chapter VII, section 5) and each f_m is a modular function of level 1, holomorphic on the upper half plane, and therefore a polynomial in $j(q)$. If we exponentiate this we see that the left hand side of 7.1 is of the form $\sum_{m\geq -1} g_m(q)p^m$ where each g_m is a polynomial in $j(q)$. Each coefficient of p^m of the right hand side of (7.1) is either a constant (if $m\neq 0$) or $744-j(q)$ (if $m=0$) and is therefore also a polynomial in $j(q)$. Any polynomial in $j(q)$ is determined by its coefficients of q^n for $n\leq 0$, so to prove the lemma it is sufficient to check that the coefficients of $q^n, n\leq 0$ of both sides are the same, which is easy to do. This proves lemma 7.1.

Remark. The infinite product in (7.1) converges if $|p|, |q| < e^{-2\pi}$ and $p \neq q$, and the infinite series for $j(q)$ converges if $|q| < 1$.

Theorem 7.2. *The simple roots of the monster Lie algebra M are the vectors $(1, n)$ ($n = -1$ or $n > 0$), each with multiplicity $c(n)$.*

Let N be the generalized Kac-Moody algebra with root lattice $II_{1,1}$ whose simple roots are as stated in the theorem. The simple roots of a generalized Kac-Moody algebra (with given Cartan subalgebra and choice of fundamental Weyl chamber) are determined by its root multiplicities because of the denominator formula, so it is sufficient to show that the multiplicity of the root (m, n) of N is equal to the multiplicity $c(mn)$ of the root (m, n) of the algebra M (so N is isomorphic to M). The denominator formula of the Lie algebra N is

$$e^{\rho} \prod_{\substack{m>0 \\ n\in\mathbf{Z}}} (1 - p^m q^n)^{\mathrm{mult}(m,n)} = \sum_{w\in W} \det(w) w(e^{\rho} S) \tag{7.2}$$

where $\mathrm{mult}(m, n)$ is the multiplicity of the root (m, n) of N, p and q are the elements $e^{(1,0)}$ and $e^{(0,1)}$ of the group ring of $II_{1,1}$, W is the Weyl group of N which is of order 2 and whose nontrivial element exchanges p and q, and e^{ρ} is $e^{(-1,0)} = p^{-1}$. By the denominator formula quoted in section 4 the element S is

$$S = \sum_{A} (-1)^{|A|} e^{\Sigma A}$$

where the sum is over all finite subsets A of the set of simple imaginary roots such that any two distinct elements of A are orthogonal, $|A|$ is the number of elements of A, and ΣA is the sum of the elements of A. All the imaginary simple roots of N have nonzero inner product with each other, so S is equal to $1 - \sum_{n>0} c(n) p q^n = 1 - p(j(q) - q^{-1} - 744)$ and hence the right hand side of the denominator formula (7.2) of N is

$$\begin{aligned} e^{\rho} S - w(e^{\rho} S) &= (p^{-1} - j(q) + q^{-1} + 744) - (q^{-1} - j(p) + p^{-1} + 744) \\ &= j(p) - j(q) \end{aligned}$$

where w is the nontrivial element of the Weyl group W. This is the right hand side of (7.1), so that the left hand sides of (7.1) and (7.2) must also be the same, which implies that the multiplicity of the root $(m, n) \neq (0, 0)$ of N is $c(mn)$. This proves theorem 7.2.

Remark. This shows that the action of the monster on M can be extended to an action of the full automorphism group of the graded vector space V on M which still preserves the grading and Lie algebra structure of M. The special property of the action of the monster on M that we use is that the pieces of degree (a, b) and (c, d) of the $II_{1,1}$-graded Lie algebra M are isomorphic as representations of the monster whenever $ab = cd \neq 0$.

8 The twisted denominator formula.

The monster Lie algebra, like any generalized Kac-Moody algebra, can be written as a direct sum $E \oplus H \oplus F$ of subalgebras, where H is the Cartan subalgebra and E and

F are the subalgebras corresponding to the positive and negative roots. We recall from section 4 the equality

$$\Lambda(E) = H(E) \tag{8.1}$$

where in the case of the monster Lie algebra both sides are virtual $II_{1,1}$-graded modules over the monster. In this section we calculate both sides of (8.1) explicitly and use this to obtain some relations between the coefficients of the Thompson series $T_g(q)$.

For any generalized Kac-Moody algebra the homology groups can be calculated as in section 4 provided we know the simple roots. In the case of the monster Lie algebra M the simple roots are given in section 7 and are just the vectors $(1, n)$, $n = -1$ or $n > 0$ with multiplicity $c(n)$, so we can calculate the homology groups $H_i(E)$ as follows. Recall that $H_i(E)$ is the subspace of elements of $\Lambda^i(E)$ whose degree r (in the root lattice) satisfies $(r, r + 2\rho) = 0$. For the monster Lie algebra we can think of the roots as elements of $II_{1,1}$ and can identify the Weyl vector ρ with the norm 0 vector $(-1, 0)$. The condition on r then says that $r + \rho$ has norm 0, i.e., r is a vector of the form $(m, 0)$ or $(1, n)$, so we just have to find the elements of $\Lambda^i(E)$ of these degrees. The answers we get are as follows (where we write $M_{m,n}$ for the subspace of elements of M of degree $(m, n) \in II_{1,1}$).

$H_0(E)$ is one dimensional with character 1, because $\Lambda^0(E) = \mathbf{R}$ is just a one dimensional space of degree $(0, 0)$.

$H_1(E)$ is the subspace of $\Lambda^1(E) = E$ of elements of degree $(m, 0)$ or $(1, n)$. There are no elements of degree $(m, 0)$ and the space of elements of degree $(1, n)$ is the simple root space $M_{1,n} \cong V_n$, so $H_1(E)$ has character $p(j(q) - 744)$ and its piece of degree $(1, n)$ is isomorphic to the n'th head representation V_n of the monster. (Notice that the first homology group $H_1(E)$ is isomorphic to the sum of the simple root spaces; the same is true for any generalized Kac-Moody algebra.)

For $i \geq 2$ there are no elements in $\Lambda^i(E)$ of degree $(1, n)$ because all elements of E have degrees of the form (m, n) with $m \geq 1$. Moreover they all have degrees with $n > 0$ except for the one dimensional space with degree equal to the real simple root $(1, -1)$, so the only way for an element of $\Lambda^i(E)$ to have degree $r = (m, 0)$ is for it to be the exterior product of two elements of E of degrees $(1, -1)$ and $(m - 1, 1)$. Therefore $H_2(E)$ is the sum of pieces of degrees $(m, 0)$ for $m \geq 2$ isomorphic to $M_{m-1,1} = V_{m-1}$, and $H_i(E) = 0$ for $i \geq 3$.

To summarize, the homology groups of E are the following $II_{1,1}$-graded vector spaces, where we use p and q to stand for one dimensional vector spaces of degrees $(1, 0)$ and $(0, 1)$ in $II_{1,1}$, and V_i has degree $0 \in II_{1,1}$ for all i.

$H_0(E) = V_{-1} = \mathbf{R}$ and has character 1.
$H_1(E) = \sum_{n\in\mathbf{Z}} V_n p q^n$ and has character $\sum_{n\in Z} c(n) p q^n = p(j(q) - 744)$.
$H_2(E) = p \sum_{m>0} V_m p^m$ and has character $p(j(p) - 744) - 1$.
$H_i(E) = 0$ if $i \geq 3$.

For most interesting generalized Kac-Moody algebras there are an infinite number of nonzero homology groups. In general the largest n with H_n nonzero is equal to (maximum length of an element of the Weyl group) + (maximum number of pairwise orthogonal imaginary simple roots) which is usually infinite; however for the monster Lie algebra both the terms in parentheses happen to be 1.

The alternating sum $\oplus(-1)^i H_i(E)$ is therefore equal to $p(\sum_m V_m p^m - \sum_n V_n q^n)$, where we use p and q to stand for one dimensional vector spaces of degrees (1,0) and (0,1). If we substitute these values of the homology groups into (8.1) we find that

$$p^{-1}\Lambda(\sum_{m>0,n\in\mathbf{Z}} V_{mn}p^m q^n) = \sum_m V_m p^m - \sum_n V_n q^n \tag{8.2}$$

where V_m is the m'th head representation of the monster simple group. (Both sides of (8.2) are essentially $II_{1,1}$-graded virtual representations of the monster simple group.) For any finite dimensional vector space U, $\Lambda(U)$ is naturally isomorphic to $\exp(-\sum_{i>0}\psi^i(U)/i)$ (where ψ^i is the i'th Adams operation) by the splitting principle, because both expressions are multiplicative in U and are equal if U is one dimensional. (The splitting principle says that two natural operations on representations are equal if they are equal on sums of 1-dimensional representations, i.e., we can pretend that any representation is a sum of 1-dimensional ones, and the Adams operation ψ^i on virtual representations of a group G is defined by $\mathrm{Tr}(g|\psi^i(U)) = \mathrm{Tr}(g^i|U)$; see [2] for more details.) The same is true for infinite dimensional vector spaces U graded by some lattice L provided that the homogeneous pieces are all finite dimensional, and the pieces of degree $a \in L$ vanish unless a is in some fixed closed cone which does not contain any line and has its vertex at the origin; this condition ensures that all the virtual vector spaces $\Lambda(U)$, $\psi^i(U)$, and so on are graded with finite dimensional pieces of each degree and are therefore meaningful. Therefore if we take the trace of some element g of the monster on both sides of (8.2) we find that

$$p^{-1}\exp(-\sum_{i>0}\sum_{m>0,n\in\mathbf{Z}}\mathrm{Tr}(g^i|V_{mn})p^{mi}q^{ni}/i) = \sum_{m\in\mathbf{Z}}\mathrm{Tr}(g|V_m)p^m - \sum_{n\in\mathbf{Z}}\mathrm{Tr}(g|V_n)q^n \tag{8.3}$$

where $\mathrm{Tr}(g|V_n)$ is the trace of g on the vector space V_n.

These relations between the coefficients of the functions $\sum \mathrm{Tr}(g|V_n)q^n$ are the same as those that Norton [28] conjectured to hold for the modular functions associated with elements of the monster in Conway and Norton [13, table 2]. The identities 8.3 imply that the functions $T_g(q)$ are completely replicable in the terminology of [28]; in fact the definition of a completely replicable function in [28] is that the function should satisfy some identities equivalent to 8.3. Norton's conjectures that these modular functions are completely replicable were proved by Koike [25]. In the next section we use this to prove that the functions $\sum_m \mathrm{Tr}(g|V_m)q^m$ are these modular functions.

9 The moonshine conjectures.

In this section we complete the proof of theorem 1.1, i.e., we verify that the monster vertex algebra V satisfies Conway and Norton's moonshine conjectures [13] that the Thompson series of elements of the monster are certain Hauptmoduls. To do this we show that the identities (8.3) in section 8 imply that the Thompson series are determined by their first 5 coefficients $c_g(i)$, $1 \leq i \leq 5$, and then use the fact that the Hauptmoduls listed by Conway and Norton in [13, table 2] satisfy the same identities ([25]) and have the same first five coefficients.

The coefficients $c_g(n) = \mathrm{Tr}(g|V_n)$ of the Thompson series of elements of the monster satisfy the relations (8.3). If we compare the coefficients of p^2 and p^4 of both sides of 8.3 and carry out some elementary algebra we find that the coefficients $c_g(i)$ satisfy the recursion formulas below for $k \geq 1$:

$$\begin{aligned}
c_g(4k) &= c_g(2k+1) + (c_g(k)^2 - c_{g^2}(k))/2 + \sum_{1\leq j<k} c_g(j)c_g(2k-j)\\
c_g(4k+1) &= c_g(2k+3) - c_g(2)c_g(2k) + (c_g(2k)^2 + c_{g^2}(2k))/2\\
&\quad + (c_g(k+1)^2 - c_{g^2}(k+1))/2 + \sum_{1\leq j\leq k} c_g(j)c_g(2k-j+2)\\
&\quad + \sum_{1\leq j<k} c_{g^2}(j)c_g(4k-4j) + \sum_{1\leq j<2k} (-1)^j c_g(j)c_g(4k-j)\\
c_g(4k+2) &= c_g(2k+2) + \sum_{1\leq j\leq k} c_g(j)c_g(2k-j+1)\\
c_g(4k+3) &= c_g(2k+4) - c_g(2)c_g(2k+1)\\
&\quad - (c_g(2k+1)^2 - c_{g^2}(2k+1))/2 + \sum_{1\leq j\leq k+1} c_g(j)c_g(2k-j+3)\\
&\quad + \sum_{1\leq j\leq k} c_{g^2}(j)c_g(4k-4j+2) + \sum_{1\leq j\leq 2k} (-1)^j c_g(j)c_g(4k-j+2)
\end{aligned} \tag{9.1}$$

where $c_g(n) = \mathrm{Tr}(g|V_n)$, $c_{g^2}(n) = \mathrm{Tr}(g^2|V_n)$. In particular if $n = 4$ or $n > 5$ then the coefficient $c_g(n)$ is determined by the coefficients $c_g(i)$ and $c_{g^2}(i)$ for $1 \leq i < n$, so if we know all the coefficients $c_g(n)$ for $n = 1$, 2, 3, and 5 and all elements g of the monster then we can work out all the coefficients $c_g(n)$. (The coefficient $c_g(5)$ is not determined by the recursion relations because they degenerate into $c_g(5) = c_g(5)$. The Thompson series of an element of the monster is not determined by its first 5 coefficients $c_g(i)$, $1 \leq i \leq 5$, but is determined by the coefficients $c_{g^j}(i)$, $1 \leq i \leq 5$ of the Thompson series of all its powers g^j. For example, the Thompson series of the elements $60F$ and $93A$ (in ATLAS [14] notation) have the same first five coefficients $c_g(i)$, $1 \leq i \leq 5$. Norton showed in [28] that any completely replicable function $\sum a(n)q^n$ is determined by its 12 coefficients a_i, $i = 1$, 2, 3, 4, 5, 7, 8, 9, 11, 17, 19, 23.)

For the case of the elliptic modular function $j(q)$, where $g = 1$, these recursion formulas were discovered by Mahler [27]. We do not make any use of the formulas (9.1), apart from the fact that the coefficients for $n = 1$, 2, 3 and 5 determine all the coefficients. (This is easy to check without calculating the formulas (9.1) explicitly. The reason for listing them explicitly is that they are a very efficient way of working out the coefficients c_g on a computer.)

Norton [28] stated and Koike [25] proved that the modular functions associated with elements of the monster in [13, table 2] also satisfy the relations (8.3) (with $\mathrm{Tr}(g|V_m)$ replaced by the coefficient of q^m of the modular function corresponding to g) and therefore also satisfy the recursion relations (9.1) above. To prove that the Thompson series of elements of the monster are these modular functions, it is therefore sufficient to check that the coefficients of q^i, $i \leq 5$ of each function are the same.

The coefficients of q^i, $i \leq 5$ of each modular function are in Conway and Norton [13, table 4], and they define representations of the monster which decompose as stated at the end of section 3. We can work out the coefficients $c_g(i)$, $1 \leq i \leq 5$ of the Thompson series of elements of the monster as follows. Knowing the coefficients $c_g(i)$, $1 \leq i \leq 5$ of the Thompson series is equivalent to knowing how the representations $V_i, 1 \leq i \leq 5$ of the monster decompose into irreducible representations of the monster. The only irreducible representations of the monster with dimension at most that of V_5 are the first seven, with characters χ_i, $1 \leq i \leq 7$ in atlas [14] notation. Therefore we can evaluate the coefficients $c_g(i)$, $1 \leq i \leq 5$ of the Thompson series of all elements of the monster provided we can find seven elements g_j, $1 \leq j \leq 7$ of the monster for which we can evaluate these coefficients, and such that the matrix $\chi_i(g_j)$, $1 \leq i, j \leq 7$ is nonsingular, because this determines the decomposition of V_i, $1 \leq i \leq 5$, into irreducible representations of the monster.

In their book [16] Frenkel, Lepowsky and Meurman give an explicit formula for the Thompson series of any element of the centralizer of an involution of type 2B of the monster. The elements of odd order of $2^{1+24}.Co_1$ correspond to the elements of odd order of the group $2.Co_1$ of automorphisms of the Leech lattice Λ. Frenkel, Lepowsky and Meurman's formula is particularly easy to evaluate if the corresponding automorphism of the Leech lattice is of odd order and fixes no nonzero vectors: it becomes

$$2\sum_n \mathrm{Tr}(g|V_n)q^n = 1/\eta_g(q) + \eta_g(q)/\eta_g(q^2) + \eta_g(q)/\eta_g(q^{1/2}) + \eta_g(q)/\eta_g(-q^{1/2}) \tag{9.2}$$

where $\eta_g(q)$ is the eta function of g acting on the Leech lattice, equal to $\eta(\epsilon_1 q)\dots\eta(\epsilon_{24}q)$ if g has eigenvalues $\epsilon_1, \dots, \epsilon_{24}$ on the vector space $\Lambda \otimes \mathbf{R}$, where $\eta(q) = q^{1/24}\prod_{n>0}(1-q^n)$. If we put $f(\tau) = 1/\eta_g(q)$ (where $q = e^{2\pi i\tau}$) then $f(\tau) - f(0)$ is the modular function of genus 0 that Conway and Norton associate to the element g of the monster in [13, table 2]. The coefficients of the right hand side of the formula above are easy to evaluate explicitly, and we can check that its coefficients of q^i for $i \leq 5$ are equal to those of the modular function $f(\tau) - f(0)$. (The right hand side of 9.2 and $f(\tau) - f(0)$ are of course equal, as we can see directly by observing that the right hand side of (9.2) is $f(\tau) + T_2 f(\tau)/f(\tau)$, where T_2 is a Hecke operator.)

For our seven elements g_j of the monster we choose an element g_1 of conjugacy classes 2B (for which Frenkel, Lepowsky and Meurman explicitly evaluate the Thompson series), and 6 more elements g_i of the monster corresponding to odd order automorphisms of the Leech lattice with no nonzero fixed vectors such that the determinant of the 7×7 matrix $\chi_i(g_j)$ is nonzero. A set of 6 elements satisfying these conditions are those of type

3B, 5B, 7B, 9B, 13B, 15D in the monster (using atlas [14] notation) corresponding to elements in the conjugacy classes 3A, 5A, 7A, 9A, 13A, and 15C of Co_1 with cycle shapes $3^{12}1^{-12}$, $5^{6}1^{-6}$, $7^{4}1^{-4}$, $9^{3}1^{-3}$, $13^{2}1^{-2}$, and $15^{2}3^{-2}$. The determinant of the matrix $\chi_i(g_j)$ is $35672555520 = 2^{22}3^{5}5^{1}7^{1}$ which is nonzero. The modular groups corresponding to these elements of the monster are $\Gamma_0(3)$, $\Gamma_0(5)$, $\Gamma_0(7)$, $\Gamma_0(9)$, $\Gamma_0(13)$, and $\binom{3,0}{0,1}^{-1}\Gamma_0(5)\binom{3,0}{0,1}$. (It is necessary to use at least one element g_i of even order, because the characters of the first seven representations of the monster are linearly dependent when restricted to elements of odd order; in fact the character of $S^2(V_1) - \Lambda^2(V_1) - V_1$ is 0 on all elements of odd order.)

We can use this to verify that the representations V_1, V_2, V_3, and V_5 of the monster decompose as stated at the end of section 3, and this implies that the numbers $\mathrm{Tr}(g|V_n)$ are equal to the coefficients of the corresponding modular functions in [13] for $n = 1$, 2, 3 and 5, and therefore for all n by the recursion relations (9.1). This completes the verification that the Thompson functions $\sum_n \mathrm{Tr}(g|V_n)q^n$ are modular functions of genus 0 and proves theorem 1.1.

10 The monstrous Lie superalgebras.

In the remainder of the paper we describe some Lie algebras similar to the monster Lie algebra. In this section we describe some associated to elements of the monster.

The twisted denominator formula for some diagram automorphism g of a generalized Kac-Moody algebra, which is obtained by taking the trace of g on both sides of the identity $H(E) = \Lambda(E)$, is often the untwisted denominator formula for some generalized Kac-Moody algebra or superalgebra. For example, the twisted denominator formula for E_6 is just the ordinary denominator formula for F_4. We construct a family of Lie algebras and superalgebras whose denominator formulas are twisted denominator formulas of the monster Lie algebra. (Warning: unlike the case of finite dimensional Lie algebras, these are not always subalgebras of the monster Lie algebra; in fact some of them are Lie superalgebras.)

If g is an element (or conjugacy class) of the monster group M then we write $T_g(q)$ for the Thompson series $T_g(q) = \sum c_g(n)q^n = \sum \mathrm{Tr}(g|V_n)q^n$ of g. For example, $T_1(q) + 744$ is the elliptic modular function $j(q)$. The function $T_g(q)$ is the normalized generator for a genus zero function field of a group containing $\Gamma_0(nh)$, where n is the order of g and h is an integer with $h|(24, n)$.

We use the convention that for super vector spaces the dimension is defined to be the dimension of the even part minus the dimension of the odd part. For example, the multiplicity of a root is the dimension of the root space, and is therefore equal to the dimension of the even root space minus the dimension of the odd root space. This provides an interpretation of roots of negative multiplicity. With this convention for multiplicities the denominator formula for generalized Kac-Moody superalgebras is the same as for generalized Kac-Moody algebras.

We define the generalized Kac-Moody superalgebra of an element g of the monster to have root lattice $II_{1,1}$ and simple roots $(1, n)$ with multiplicity $\mathrm{Tr}(g|V_n)$. (These multiplicities are all integers as can be seen in several ways; for example they are rational by the recursion relations 9.1 and algebraic integers because they are character values.) This is really a family of Lie superalgebras rather than a single superalgebra, because for each simple root $(1,n)$ we can choose any two nonnegative integers a and b with $a-b=\mathrm{Tr}(g|V_n)$ and take a even simple roots $(1,n)$ and b odd ones. This only seems to be interesting if all coefficients of T_g are positive, in which case we get a Lie algebra, or if the coefficients alternate in sign, in which case we get a superalgebra such that the super elements are those of degree (m,n) with $m+n$ odd.

We can work out the root multiplicities of these superalgebras because their denominator formulas are the twisted denominator formulas of the monster Lie algebra. The denominator formula for the Lie superalgebra of g is

$$p^{-1}\prod_{m>0,n\in\mathbf{Z}}(1-p^mq^n)^{\mathrm{mult}(m,n)}=\sum_m \mathrm{Tr}(g|V_m)p^m-\sum_n \mathrm{Tr}(g|V_n)q^n=T_g(p)-T_g(q) \quad (10.1)$$

and by (8.3)

$$T_g(p)-T_g(q)=p^{-1}\exp(-\sum_{i>0}\sum_{m>0}\sum_{n\in\mathbf{Z}}\mathrm{Tr}(g^i|V_{mn})p^{mi}q^{ni}/i). \quad (10.2)$$

If we compare the logarithms of (10.1) and (10.2) we find that

$$\sum_{i>0,m>0,n\in\mathbf{Z}}\mathrm{Tr}(g^i|V_{mn})p^{mi}q^{ni}/i=\sum_{i>0,m>0,n\in\mathbf{Z}}\mathrm{mult}(m,n)p^{mi}q^{ni}/i \quad (10.3)$$

and applying the Moebius inversion formula to (10.3) shows that

$$\mathrm{mult}(m,n)=\sum_{ds|(m,n)}\frac{\mu(s)}{ds}\mathrm{Tr}(g^d|V_{mn/d^2s^2}) \quad (10.4)$$

where μ is the Moebius function. If f is any function of n whose value depends only on (n,N) for some N, then the sum $\sum_{ds=k}\mu(s)f(d)$ is 0 unless $k|N$. The trace $\mathrm{Tr}(g^d|V_n)$ depends only on n and (d,N) where N is the order of g, so $\sum_{ds=k}\mu(s)\mathrm{Tr}(g^d|V_{mn/d^2s^2})$ is 0 unless $k|N$, hence (10.4) can be simplified to

$$\mathrm{mult}(m,n)=\sum_{ds|(m,n,N)}\frac{\mu(s)}{ds}\mathrm{Tr}(g^d|V_{mn/d^2s^2}) \quad (10.5)$$

This is the formula for the multiplicity of the roots of the superalgebra of g. The sum on the right of (10.5) can often be simplified by expressing T_{g^d} in terms of T_g. For example, suppose that $N=N_1N_2$ is a squarefree integer such that the group $\Gamma_0(N)+d|N_1$ has genus 0, where $\Gamma_0(N)+d|N_1$ is the group generated by $\Gamma_0(N)=\{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in SL_2(\mathbf{Z})|c\equiv 0 \bmod N\}$ together with the Atkin-Lehner involutions of all divisors of N_1. (See [13].) If g is an element of the monster corresponding to one of these groups then so is g^k. By the compression formula of [13] section 8

$$T_{g^p}(\tau)=T_g(\tau)+T_g(\tau/p)+\ldots+T_g((\tau+p-1)/p)$$

for any prime divisor p of N_1, and

$$T_{g^p}(\tau) = T_g(\tau) - T_g(\tau/p^2) - \ldots - T_g((\tau + p^2 - 1)/p^2)$$

for any prime divisor p of N_2. These imply that

$$\mathrm{Tr}(g^d|V_n) = \sum_{d_1|(d,N_1)} \sum_{d_2|(d,N_2)} d_1 d_2^2 \mu(d_2) \mathrm{Tr}(g|V_{d_1 d_2^2 n}) \tag{10.6}$$

by induction on the number of divisors of d. If we substitute (10.6) into (10.5) and rearrange we find that

$$\mathrm{mult}(m,n) = \sum_{d|(m,n,N_1)} \mathrm{Tr}(g|V_{mn/d}) \prod_{p|(m,n,N_2)} (1-p) \tag{10.7}$$

where the product is over all primes p dividing (m, n, N_2).

We get particularly interesting results when N_2 is 1 or 2. First suppose that $N_2 = 1$ so that $N = N_1$ is one of the 44 squarefree integers that are orders of elements in the monster, which are the same as the integers for which $\Gamma_0(N)+$ has genus 0. Let g be an element of the monster corresponding to the group $\Gamma_0(N)+$, so that $T_g(q)$ is the normalized generator of the function field of $\Gamma_0(N)+$. Then (10.7) becomes

$$\mathrm{mult}(m,n) = \sum_{d|(m,n,N)} \mathrm{Tr}(g|V_{mn/d}) \tag{10.8}$$

and if we substitute (10.8) into (10.1) we obtain the formula

$$T_g(p) - T_g(q) = p^{-1} \prod_{m>0, n\in\mathbf{Z}} \prod_{d|N} (1 - p^{dm} q^{dn})^{c_g(dmn)}$$

which generalizes the product formula for the j function. For these cases the coefficients c_g of the Thompson series $T_g(q)$ are always nonnegative, so the algebras we get are Lie algebras rather than Lie superalgebras. This seems to happen whenever the corresponding subgroup of $PGL_2(R)$ contains the Fricke involution $\tau \to -1/N\tau$. More generally, Norton has observed ([29]) that the result of applying the Fricke involution to a Thompson series always has nonnegative coefficients.

Similarly suppose that $N_2 = 2$ so that N_1 is one of the squarefree integers 1, 3, 5, 7, 11, or 15 for which $\Gamma_0(2N_1) + d|N_1$ has genus 0. Then (10.7) becomes

$$\mathrm{mult}(m,n) = (-1)^{(m-1)(n-1)} \sum_{d|(m,n,N_1)} \mathrm{Tr}(g|V_{mn/d}).$$

The coefficients of the Thompson series for these six cases alternate in sign, so the multiplicity of (m, n) is at least 0 or at most 0 depending on whether $m + n$ is even or odd. Therefore the root (m, n) is an even root or an odd (super) root depending on the parity of $m + n$. The same is true for any element whose Thompson series has coefficients that alternate in sign. If the coefficients of a Thompson series are neither at least 0 nor alternating in sign, then the root space of some root (m, n) is in general a sum of an even space and an odd space, and the multiplicity of the root is the difference of the dimensions of these two spaces. However the Lie superalgebras are probably not very interesting unless the coefficients are either all positive or alternating in sign.

11 Some modular forms.

We verify some identities involving modular forms of level 2 that we use later. This section can be missed out if the reader is willing to assume these identities.

The modular forms in this section are usually forms of level 2, i.e. modular forms for the group $\Gamma(2) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbf{Z}) | \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2\}$. We recall a few facts about this group from [22]. The group $\Gamma(2)$ has no elliptic elements, 3 cusps (represented by 0, 1, and ∞), genus 0, and has index 6 in $SL_2(\mathbf{Z})$. The dimension of the space of forms of even nonnegative weight $2k$ is equal to $k+1$ and if f and g are two linearly independent forms of weight 2 then every form can be written uniquely as a polynomial in f and g. Two modular forms of level 2 and weight $2k$ are equal if and only if their coefficients of q^0, $q^{1/2}$, q^1, ... , q^k are equal. The group $\Gamma(2)$ is conjugate to the group $\Gamma_0(4) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbf{Z}) | c \equiv 0 \bmod 4\}$, and $f(q)$ is a modular form for $\Gamma_0(4)$ if and only if $f(q^{1/2})$ is a modular form for $\Gamma(2)$.

We list several modular forms of level 2, together with enough of their Fourier coefficients to verify the identities we use in later sections. Recall that to prove that two modular forms are equal it is sufficient to check that sufficiently many of their Fourier coefficients are equal.

The following are some forms of weight 4 and level 2.

$$\begin{aligned}
\theta_{E_8}(q) &= 1 + 240(q + 9q^2 + 28q^3 + \ldots + \sigma_3(n)q^n + \ldots) \\
\theta_{2E_8}(q) &= 1 + 240q^2 + \ldots \\
\theta_{2E_8+v_1}(q) &= 2q^{1/2} + 56q^{3/2} + 252q^{5/2} + \ldots \\
\theta_{2E_8+v_2}(q) &= 16q + 128q^2 + 448q^3 + \ldots \\
\eta(q)^{16}\eta(q^{1/2})^{-8} &= q^{1/2} + 8q + 28q^{3/2} + 64q^2 + 126q^{5/2} + 224q^3 + \ldots \\
\eta(q)^{16}\eta(-q^{1/2})^{-8} &= -q^{1/2} + 8q - \ldots \\
\eta(q)^{16}\eta(q^2)^{-8} &= 1 - 16q + 112q^2 - 448q^3 + \ldots
\end{aligned}$$

The form $\theta_{E_8}(q)$ is the theta function of the E_8 lattice, and the form $\theta_{2E_8+v_i}$ is the theta function of the coset of $2E_8$ containing a vector v_i of norm i in the dual of $2E_8$ but not in E_8 (where $2E_8$ is the E_8 lattice with all norms doubled).

Lemma 11. 1. *If* $\eta_{2+}(q) = \eta(q)^8\eta(q^2)^8$ *then*

$$\begin{aligned}
\theta_{2E_8}(q)/\Delta(q) &= 1/\eta_{2+}(q) + 1/\eta_{2+}(q^{1/2}) + 1/\eta_{2+}(-q^{1/2}) \\
\theta_{2E_8+v_1}(q)/\Delta(q) &= 1/\eta_{2+}(q^{1/2}) - 1/\eta_{2+}(-q^{1/2}) \\
\theta_{2E_8+v_2}(q)/\Delta(q) &= 1/\eta_{2+}(q^{1/2}) + 1/\eta_{2+}(-q^{1/2})
\end{aligned}$$

Proof: If we multiply both sides by $\Delta(q)$ then both sides are modular forms of level 2 and weight 4, so it is sufficient to check that the coefficients of q^0, $q^{1/2}$, and q^1 of both sides are equal, which we can do using the expressions above.

The following are some forms of weight 8 and level 2.

$$\begin{aligned}
\theta_{\Lambda_2}(q) &= 1 + 4320q^2 + \dots \\
\theta_{\Lambda_2}(q^{1/2}) &= 1 + 4320q^1 + 61440q^{3/2} + 522720q^2 + \dots \\
\theta_{\Lambda_2}(-q^{1/2}) &= 1 + 4320q - 61440q^{3/2} + 522720q^2 + \dots \\
\eta(q)^{32}\eta(q^2)^{-16} &= 1 - 32q + 480q^2 - 4480q^3 + \dots \\
\eta(q)^{32}\eta(q^{1/2})^{-16} &= q + 16q^{3/2} + 120q^2 + 576q^{5/2} + 2060q^3 + \dots \\
\eta(q)^{32}\eta(-q^{1/2})^{-16} &= q - 16q^{3/2} + 120q^2 \dots
\end{aligned}$$

The form $\theta_{\Lambda_2}(q)$ is the theta function of the Barnes-Wall lattice Λ_2 of dimension 16, and is a modular form for $\Gamma_0(2)+$. The forms $\theta_{\Lambda_2+v_i}$ are defined in the same way as for the lattice $2E_8$.

Lemma 11. 2. *Let $\eta_{2-}(q)$ be the modular form $\eta(q)^{-8}\eta(q^2)^{16}$. Then*

$$\begin{aligned}
\theta_{\Lambda_2}(q)/\Delta(q) &= 1/\eta_{2-}(q) - 1/\eta_{2-}(q^{1/4}) - 1/\eta_{2-}(iq^{1/4}) - 1/\eta_{2-}(-q^{1/4}) - 1/\eta_{2-}(-iq^{1/4}) \\
\theta_{\Lambda_2+v_3}/\Delta(q) &= 1/\eta_{2-}(q^{1/4}) + 1/\eta_{2-}(iq^{1/4}) + 1/\eta_{2-}(-q^{1/4}) + 1/\eta_{2-}(-iq^{1/4}) \\
\theta_{\Lambda_2+v_2}/\Delta(q) &= 1/\eta_{2-}(q^{1/4}) - 1/\eta_{2-}(iq^{1/4}) + 1/\eta_{2-}(-q^{1/4}) - 1/\eta_{2-}(-iq^{1/4})
\end{aligned}$$

Proof. If we multiply both sides by $\Delta(q)$ then both sides are modular forms of weight 8 and level 2, so it is sufficient to check that the coefficients of q^0, $q^{1/2}$, q^1, $q^{3/2}$, and q^2 of both sides are equal.

12 The fake monster Lie algebra.

In the next section we construct a family of superalgebras whose denominator formulas are twisted versions of the denominator formula of the fake monster Lie algebra, in the same way that we constructed a family of algebras from the monster Lie algebra. To do this we need a detailed description of the fake monster Lie algebra, which we give in this section.

As the Leech lattice Λ is an even lattice, it has a unique central extension

$$0 \to \{1, -1\} \to \hat{\Lambda} \to \Lambda \to 0$$

by a group of order 2, such that the commutator of any inverse images of $r, s \in \Lambda$ is $(-1)^{(r,s)}$. This central extension is unique up to nonunique isomorphism, and its group of automorphisms $\mathrm{Aut}(\hat{\Lambda})$ preserving the inner product is a nonsplit extension $2^{24}.\mathrm{Aut}(\Lambda)$. (The double cover of this nonsplit extension is isomorphic to the double cover of a centralizer of an involution of type 2B in the monster simple group.)

We start by recalling some results about the fake monster Lie algebra from [8] (where it is called the monster Lie algebra). The root lattice of the fake monster Lie algebra is the 26 dimensional even unimodular Lorentzian lattice $II_{25,1} = \Lambda \oplus II_{1,1}$, where the norm of an element $(\lambda, m, n) \in \Lambda \oplus II_{1,1}$ is defined to be $\lambda^2 - 2mn$. The real simple roots are the norm 2 vectors of the form $(\lambda, 1, \lambda^2/2-1)$, and the imaginary simple roots are the positive multiples $n\rho$ of the Weyl vector $\rho = (0, 0, 1)$, each with multiplicity 24. A nonzero vector $r \in II_{25,1}$ is a root if and only if $r^2 \leq 2$, in which case it has multiplicity $p_{24}(1 - r^2/2)$ where $p_{24}(n)$ is the number of partitions of n into parts of 24 colours. The fake monster Lie algebra is essentially the space of physical vectors of the vertex algebra of $\hat{II}_{25,1}$, and can be constructed from the vertex algebra V_Λ of $\hat\Lambda$ in the same way that the monster Lie algebra M was constructed from the monster vertex algebra V in section 6.

The fake monster Lie algebra is acted on by the group of affine automorphisms of the Leech lattice which is a split extension $\Lambda.\mathrm{Aut}(\Lambda)$, because its Dynkin diagram may be interpreted as the Leech lattice as it has one real simple root $(\lambda, 1, \lambda^2/2 - 1)$ for each Leech lattice vector λ. However this is the wrong group of automorphisms to use. A more useful group of automorphisms of the fake monster Lie algebra is the nonsplit extension $2^{24}.\mathrm{Aut}(\Lambda) = \mathrm{Aut}(\hat\Lambda)$, which acts naturally on the vertex algebra of the central extension $\hat\Lambda$ of the Leech lattice Λ, and therefore on the fake monster Lie algebra. The point is that we can use the no-ghost theorem 5.1 to describe how the group $\mathrm{Aut}(\hat\Lambda)$ acts on the root spaces of the fake monster Lie algebra, and it is not obvious how we could do this if we used the group $\mathrm{Aut}(\Lambda)$.

We now describe what the fake monster Lie algebra M_Λ looks like as a module over $\mathrm{Aut}(\hat\Lambda) = 2^{24}\mathrm{Aut}(\Lambda)$. If we forget the Λ in the $II_{25,1} = \Lambda \oplus II_{1,1}$ grading of M_Λ then we can consider M_Λ to be a $II_{1,1}$-graded Lie algebra. The piece of M_Λ of degree $(m, n) \in II_{1,1}$ is then isomorphic as a Λ-graded $2^{24}\mathrm{Aut}(\Lambda)$ module to the weight $1 - mn$ piece of the vertex algebra V_Λ of $\hat\Lambda$ by theorem 6.1, so it is sufficient to describe V_Λ explicitly, which we can do as follows.

We let $S = S(\oplus_{i>0}\Lambda_i)$ be the symmetric algebra on the sum of a countable number of copies Λ_i of the Leech lattice Λ, and grade this by letting the elements of Λ_i have weight i. The dimension of the subspace of S of weight n is then equal to $p_{24}(n)$, the number of partitions of n into parts of $24 = \dim(\Lambda)$ colours, which is the coefficient of q^{n-1} of $1/\Delta(q) = q^{-1}\prod_{i>0}(1 - q^i)^{-24}$. We let $\mathbf{Z}(\hat\Lambda)$ be the twisted group ring of Λ, which is the group ring of $\hat\Lambda$ quotiented out by the ideal $(1 + \epsilon)$ where ϵ is the element of the group ring of $\hat\Lambda$ corresponding to $-1 \in \hat\Lambda$, and grade it by letting e^v have weight $(v, v)/2$ for $v \in \hat\Lambda$. (This grading does not have the property that $\deg(ab) = \deg(a) + \deg(b)$.) Finally the graded space V_Λ is isomorphic to $\mathbf{Z}(\hat\Lambda) \otimes S$. The group $2^{24}\mathrm{Aut}(\Lambda) = \mathrm{Aut}(\hat\Lambda)$ acts naturally on both $\hat\Lambda$ and on Λ_i, and so acts on V_Λ. Finally the Λ grading on V_Λ comes from the Λ grading on $\mathbf{Z}(\hat\Lambda)$. (The Λ-grading on S is the trivial one such that everything has degree 0.) This gives a complete description of V_Λ as a Λ-graded $\mathrm{Aut}(\hat\Lambda)$ module, and therefore describes M_Λ as a $II_{25,1}$-graded $\mathrm{Aut}(\hat\Lambda)$ module.

We need a lemma to construct some elements of $\mathrm{Aut}(\hat\Lambda)$ from elements of $\mathrm{Aut}(\Lambda)$.

Lemma 12.1. *Let σ be an automorphism of Λ of order n, let Λ^σ the vectors of Λ fixed by σ, and let $\hat\Lambda^\sigma$ be the inverse image of Λ^σ in $\hat\Lambda$. If n is odd then σ can be lifted to an element $\hat\sigma \in \mathrm{Aut}(\Lambda)$ of order n which fixes all elements of $\hat\Lambda^\sigma$. If n is even then σ can be lifted to an element $\hat\sigma \in \mathrm{Aut}(\Lambda)$ which fixes all elements of $\hat\Lambda^\sigma$. The automorphism $\hat\sigma^n$ multiplies any element v of $\hat\Lambda$ by $\epsilon^{(v,\sigma^{n/2}(v))}$, and in particular $\hat\sigma$ has order n if $(v, \sigma^{n/2}(v))$ is even for all v, and has order $2n$ otherwise.*

Proof. If n is odd this result is obvious because σ has lifts of the same odd order, and any lift of odd order must fix all elements of $\hat{\Lambda}^g$. If n is even then we can multiply any lift of σ by some element of 2^{24} to get an element $\hat{\sigma}$ which fixes all elements of $\hat{\Lambda}^\sigma$. The element $v\sigma(v)\ldots\sigma^{n-1}(v)$ is fixed by σ as it is in $\hat{\Lambda}^\sigma$ and is therefore equal to $\sigma(v)\ldots\sigma^n(v)$. Hence $v=\sigma^n(v)$ if and only if v commutes with $\sigma(v)\ldots\sigma^{n-1}(v)$, which is true if and only if $(v,\sigma(v))+\ldots(v,\sigma^{n-1}(v))$ is even, which is true if and only if $(v,\sigma^{n/2}v)$ is even because $(v,\sigma^i(v))=(v,\sigma^{n-i}(v))$. Therefore $\sigma^n(v)$ fixes v if and only if $(v,\sigma^{n/2}(v))$ is even, and therefore multiplies v by $\epsilon^{(v,\sigma^{n/2}(v))}$. This proves lemma 12.1.

For example, if σ is an involution of $\mathrm{Aut}(\Lambda)$ whose lattice Λ^σ has dimension 8 or 16 then it is easy to check that $(v,\sigma(v))$ is even for all $v\in\Lambda$, so σ can be lifted to an element $\hat{\sigma}\in\mathrm{Aut}(\hat{\Lambda})$ which has order 2 and fixes all elements of $\hat{\Lambda}^\sigma$. This is not possible if σ is an involution with Λ^σ 12-dimensional, because there are vectors v such that $(v,\sigma(v))$ is odd.

The homology $H(E)$ of the subalgebra E of M_Λ corresponding to the positive roots can be described as follows. We let e^ρ stand for a one dimensional $II_{25,1}$-graded vector space of degree $\rho\in II_{25,1}$. Then the virtual vector space $e^\rho H(E)$ is antisymmetric under the Weyl group, so it is only necessary to describe the subspace of it generated by elements whose degrees are in the Weyl chamber of $II_{25,1}$. By the remarks at the end of section 4 this subspace is naturally isomorphic to the virtual vector space $\Lambda(\oplus_{i>0}\Lambda_i)$, where Λ_i is a copy of the vector space of Λ with degree $i\rho\in II_{25,1}$. This follows from the fact that the imaginary simple roots of M_Λ are the positive multiples of ρ each with multiplicity 24, so all imaginary simple roots are orthogonal to each other, so the subspace of the homology group that we need is just the exterior power of the sum of the root spaces of the imaginary simple roots.

13 The denominator formula for fake monster Lie algebras.

We can now calculate a twisted denominator formula for each element of $\mathrm{Aut}(\hat{\Lambda})$, which is often the denominator formula for some other Lie algebra or superalgebra. In this section we calculate a few examples of these Lie algebras.

We let g be some element of $\mathrm{Aut}(\hat{\Lambda})=2^{24}\mathrm{Aut}(\Lambda)$ of order N, and we let L be the sublattice of Λ fixed by g. The projection of Λ into the vector space of L is the dual lattice L' of L because Λ is unimodular. For simplicity we assume that any power g^n of g fixes all elements of $\hat{\Lambda}$ which are in the inverse image of any vector of Λ fixed by g^n. (If we do not assume this condition we can still carry out most of the work here, but we would have to put more effort into keeping track of signs.) Lemma 12.1 provides us with plenty of examples of such elements g; for example any element of odd order.

We consider both sides of the formula

$$\Lambda(E)=H(E)$$

to be L'-graded virtual g modules. We calculate the trace of g on both sides, which can be thought of as an element of some completion of the group ring of L'. We can calculate the trace of g on $\Lambda(E)$ provided we know exactly how g acts on E, and we can calculate the trace of g on $H(E)$ if we know exactly how g acts on the simple root spaces of M_Λ.

We first calculate the trace of g on $\Lambda(E) = \exp(-\sum_{i>0}\psi^i(E)/i)$ (where ψ^i are the Adams operations). This trace is equal to

$$\exp(-\sum_{v\in L'}\sum_{i>0}\mathrm{Tr}(g^i|E_v)e^{iv}/i) \tag{13.1}$$

where E_v is the subspace of E whose L'-degree is v. (Note added May 1999: the published version of this formula incorrectly had a term E_{iv} instead of E_v.) We wish to express the trace of g in the form $\prod(1-e^r)^{\mathrm{mult}(r)} = \exp - \sum_r\sum_{i>0}e^{ir}\mathrm{mult}(r)/i$ for some numbers $\mathrm{mult}(r)$. By applying the Moebius inversion formula to (13.1) as in section 11 we see that the numbers $\mathrm{mult}(r)$ are given by

$$\mathrm{mult}(r) = \sum_{ds|((r,L),N)}\mu(s)\mathrm{Tr}(g^d|E_{r/ds})/ds \tag{13.2}$$

where (r, L) is the highest common factor of the numbers (r, a) for $a \in L$.

By theorem 2.2 of [6] there is a reflection group W^g acting on the lattice L, which may be taken as any of the following groups.

1 The subgroup of W of elements that commute with g.

2 The subgroup of W of elements that map ρ into L.

3 The subgroup of W of elements mapping L into L.

The positive roots of W^g are the vectors which are the sums of the conjugates of some positive real root of $II_{25,1}$. The simple roots of W^g are the sums of orbits of simple roots of W that have positive norm, and they are also the roots of W^g satisfying $(r,\rho) = -r^2/2$, so ρ is a norm 0 Weyl vector for W^g. (Warning: W^g is not always the full reflection group of L.)

From (13.2) we can see that any 2 positive vectors of L' conjugate under W^g have the same multiplicity. If r is a real simple root of W^g then an easy calculation using (13.2) shows that $\mathrm{mult}(r)$ is 1, so $\mathrm{mult}(r)$ is 1 for any positive real root of W^g, because any such root is conjugate under W^g to a simple root of W^g. If r has positive norm but is not a root of W^g then $\mathrm{mult}(r) = 0$.

The product

$$e^\rho\prod_{r\in L}(1-e^r)^{\mathrm{mult}(r)}$$

is therefore antisymmetric under the reflection group W^g, so it is equal to $\sum_{w\in W^g}\det(w)w(e^\rho S)$ for some S with $e^\rho S$ in the Weyl chamber of W^g. (The determinant of w means its determinant as an automorphism of L, rather than as an automorphism of Λ.) This sum is also equal to the trace of g on $e^\rho H(E)$ because it is equal to the trace of g on $e^\rho\Lambda(E)$, so S is the trace of g on the subspace of $H(E)$ in the Weyl chamber of W^g. This subspace is isomorphic to $\Lambda(\oplus_{i>0}\Lambda_i)$ where Λ_i is a copy of the Leech lattice with degree $i\rho \in II_{25,1}$. The trace of g on this subspace is just $e^{-\rho}\eta_g(e^\rho)$ where $\eta_g(q)$ is defined to be $\eta(\epsilon_1 q)\ldots\eta(\epsilon_{24}q)$ if g has eigenvalues ϵ_1, ... , ϵ_{24} on the vector space of Λ. (If G has generalized cycle shape $a_1^{b_1}a_2^{b_2}\ldots$ then $\eta_g(q) = \eta(q^{a_1})^{b_1}\eta(q^{a_2})^{b_2}\ldots$.) Therefore we finally get

$$e^\rho\prod_{r\in L}(1-e^r)^{\mathrm{mult}(r)} = \sum_{w\in W^g}\det(w)w(\eta_g(e^\rho)). \tag{13.3}$$

The right hand side of this is exactly the denominator formula for a generalized Kac-Moody superalgebra with the following simple roots.

1 The real simple roots are the simple roots of the reflection group W^g, which are the roots r with $(r,\rho) = -(r,r)/2$.

2 The imaginary simple roots are the positive multiples $i\rho$ of the Weyl vector ρ, with multiplicity equal to

$$\mathrm{mult}(i\rho) = \sum_{ja_k=i} b_k$$

if g has generalized cycle shape $a_1^{b_1}a_2^{b_2}\ldots$.

Therefore the left hand side is the left hand side of the denominator formula for this generalized Kac-Moody superalgebra, so its positive roots have multiplicity $\mathrm{mult}(r)$.

14 Examples of fake monster Lie algebras.

We now evaluate both sides of the denominator formula (13.3) for several elements of $\mathrm{Aut}(\hat\Lambda)$, so we get several example of generalized Kac-Moody algebras whose simple roots and root multiplicities are known explicitly.

Example 1. The fake baby monster Lie algebra of rank 18. For this example we let $g \in \mathrm{Aut}(\hat\Lambda)$ be an element of order 2 which is the lift of an element of order 2 of $\mathrm{Aut}(\Lambda)$ which fixes a 16 dimensional sublattice Λ^g of Λ, such that g fixes all elements of $\hat\Lambda^g$. By lemma 12.1 such an element g exists, and the lattice Λ^g is the Barnes-Wall lattice of dimension 16 and determinant 2^8 [12]. The root lattice L is the sum of Λ^g and the two dimensional even Lorentzian lattice $II_{1,1}$. The Barnes-Wall lattice Λ^g has no roots, so by theorem 3.1 of [6] the group W^g is the full reflection group of L.

We can calculate the multiplicities $\mathrm{mult}(r)$ explicitly. The result we get is that $\mathrm{mult}(r) = p_g(1-r^2/2)$ if $r \in L$ and $r \notin 2L'$, $\mathrm{mult}(r) = p_g(1-r^2/2) + p_g(1-r^2/4)$ if $r \in 2L'$, and $\mathrm{mult}(r) = 0$ otherwise, where $p_g(1+n)$ is the coefficient of q^n in $1/\eta_g(q) = q^{-1}\prod_{i>0}(1-q^i)^{-8}(1-q^{2i})^{-8} = q^{-1} + 8 + 44q + \ldots$. If $r \in L'$ then $\mathrm{Tr}(g|E_r)$ is 0 if $r \notin L$ and equal to the coefficient of $q^{-r^2/2}$ in $1/\eta_g(q)$ if $r \in L$. The value of $\mathrm{Tr}(g^2|E_r)$ is just the dimension of E_r, which is the coefficient of $q^{-r^2/2}$ in $\theta_{\Lambda^{g\perp}+r^\perp}(q)/\Delta(q)$ where $\theta_{\Lambda^{g\perp}+r^\perp}(q)$ is the theta function of the coset $\Lambda^{g\perp}+r^\perp$ of the lattice $\Lambda^{g\perp}$ and where $r^\perp$ is a vector such that $r + r^\perp$ is in $II_{25,1}$. If we substitute these values into the formula

$$\mathrm{mult}(r) = \mathrm{Tr}(g|E_r) - \mathrm{Tr}(g|E_{r/2})/2 + \mathrm{Tr}(g^2|E_{r/2})/2$$

we find that $\mathrm{mult}(r)$ is equal to the coefficient of $q^{-r^2/2}$ in

$$\begin{aligned} 0 &\text{ if } r \notin L\\ 1/\eta_g(q) &\text{ if } r \in L, r \notin 2L'\\ 1/\eta_g(q) + \theta_{\Lambda^{g\perp}+r^\perp}(q^4)/\Delta(q^4) &\text{ if } r \in 2L', r \notin 2L\\ 1/\eta_g(q) - 1/\eta_g(q^4)/2 + \theta_{\Lambda^{g\perp}}(q^4)/\Delta(q^4)/2 &\text{ if } r \in 2L. \end{aligned}$$

By lemma 11.1 this implies that if $r \in 2L'$ then $\mathrm{mult}(r)$ is equal to the coefficient of $q^{-r^2/2}$ of $1/\eta_g(q) + 1/\eta_g(q^2)$. Therefore the explicit version of the product formula for the fake baby monster Lie algebra is

$$e^\rho \prod_{r\in L^+}(1-e^r)^{p_g(1-r^2/2)}\prod_{r\in 2L'^+}(1-e^r)^{p_g(1-r^2/4)} = \sum_{w\in W^g}\det(w)w(e^\rho\prod_{i>0}(1-e^{i\rho})^8(1-e^{2i\rho})^8)$$

where L is the Lorentzian lattice which is the sum of the Barnes-Wall lattice Λ^g and the two dimensional even Lorentzian lattice, W^g is its reflection group which has Weyl vector ρ, and L' is the dual of L.

Example 2. We can construct a similar Lie algebra for any of the primes $p = 2$, 3, 5, 7, 11, or 23 with $(p+1)|24$. We let g be an element of $\mathrm{Aut}(\hat\Lambda)$ of order p corresponding to an element of $M_{24}\subset \mathrm{Aut}(\Lambda)$ of cycle shape $1^{24/(p+1)}p^{24/(p+1)}$. We assume that the lattice Λ^g of dimension $48/(p+1)$ has no roots and if $r\in\Lambda^{g\perp\prime}$ then

$$\theta_{\Lambda^{g\perp}+r}(q) = \eta(q)^{24p/(p+1)}(\eta(q^p)^{-24/(p+1)}\delta(r\in\Lambda^{g\perp}) + \sum_{0\le i<p}\epsilon^{-ipr^2/2}\eta(\epsilon^i q^{1/p})^{-24/(p+1)})$$

where ϵ is a primitive p'th root of 1 and where $\delta(r\in\Lambda^{g\perp})$ is 1 if $r\in\Lambda^{g\perp}$ and 0 otherwise. These assumptions are in principle not difficult to prove if they are true, as they are essentially just identities between modular forms. (For $p = 2$ they are just lemma 12.1.)

Assuming these identities, an argument similar to the one above for $p = 2$ shows that the denominator formula for the Lie algebra of G is

$$\begin{aligned} &e^\rho\prod_{r\in L^+}(1-e^r)^{p_g(1-r^2/2)}\prod_{r\in pL'^+}(1-e^r)^{p_g(1-r^2/2p)}\\ =&\sum_{w\in W^g}\det(w)w(e^\rho\prod_{i>0}(1-e^{i\rho})^{24/(p+1)}(1-e^{pi\rho})^{24/(p+1)}) \end{aligned}$$

where $\sum_{i>0}p_g(1+i)q^i = 1/\eta_g(q)$. For p= 2, 3, 5, 7, and 11 these Lie algebras seem to correspond to the baby monster, the Fischer group Fi_{24}, the Harada Norton group, the Held group and the Mathieu group M_{12} in the same way that the fake monster Lie algebra corresponds to the monster simple group.

Example 3. The fake Conway Lie superalgebra of rank 10. This is the algebra described at the end of section 2, and seems to correspond to Conway's simple group Co_1. We let $g\in\mathrm{Aut}(\hat\Lambda)$ be an element of order 2 which is the lift of an element of order 2 of $\mathrm{Aut}(\Lambda)$ which fixes a lattice Λ^g of Λ of dimension 8. This lattice Λ^g is isomorphic to the E_8 lattice with all norms doubled, so if we halve all the norms of the lattice $\Lambda^g\oplus II_{1,1}$ we get the nonintegral lattice of determinant 1/4 which is the dual of the sublattice of even vectors of $I_{9,1}$. A calculation similar to that in example 1 but using lemma 11.2 shows that the denominator formula of this Lie superalgebra is

$$e^\rho\prod_{r\in\Pi^+}(1-e^r)^{\mathrm{mult}(r)} = \sum_{w\in W}\det(w)w(e^\rho\prod_{n>0}(1-e^{n\rho})^{(-1)^n 8})$$

where the multiplicity of the root $r = (v, m, n)\in L$ is equal to

$$\mathrm{mult}(r) = (-1)^{(m-1)(n-1)}p_g((1-r^2)/2) = (-1)^{m+n}|p_g((1-r^2)/2)|,$$

and $p_g(n)$ is defined by

$$\sum p_g(n)q^n = q^{-1/2}\prod_{n>0}(1-q^{n/2})^{-(-1)^n 8}.$$

15 Open questions

We list a few conjectures and open questions about the Lie algebras and superalgebras we have constructed.

(1) Prove the assumptions about the modular forms used in example 2 in section 14. In principle this should be easy (at least when they are true) because they just involve checking a finite number of identities between modular forms.

(2) Investigate the Lie algebras and superalgebras coming from other elements of the monster or $\mathrm{Aut}(\Lambda)$ and write down their denominator formulas explicitly in some nice form.

(3) Find a natural construction for these Lie algebras and superalgebras (i.e. other than by generators and relations). We used natural constructions for the monster Lie algebra and the fake monster Lie algebra from vertex algebras, but I do not know of any similar constructions for most of the other Lie algebras. The easiest case is the superalgebra of rank 10 which can be constructed from superstrings on a 10-dimensional torus. (The even part of the rank 10 superalgebra was constructed in [3, 4]; the odd part is more difficult to construct.) A natural construction should give actions of various finite groups on these Lie algebras; for example the double cover of the baby monster should act on the baby monster Lie algebra.

(4) Describe the Lie bracket from $V_{ab}\otimes V_{cd}$ to $V_{(a+c)(b+d)}$ of the monster Lie algebra explicitly in terms of the vertex algebra operations on V.

(5) Is the baby monster Lie algebra a subalgebra of the monster Lie algebra in a way that preserves the action of the double cover of the baby monster? Similarly for the other monstrous Lie algebras.

(6) Are there any generalized Kac-Moody algebras, other than the finite dimensional, affine, monstrous or fake monstrous ones, whose simple roots and root multiplicities can both be described explicitly? The monstrous and fake monstrous algebras are both finite families, each with a few hundred members, corresponding roughly to the conjugacy classes in the monster and in $\mathrm{Aut}(\hat{\Lambda})$. One example of a denominator formula for a Lie superalgebra of rank 3 is the identity

$$\sum_{i+j+k=0}(-p)^{jk}(-q)^{ik}(-r)^{ij} = \prod_{i+j+k>0}((1-p^iq^jr^k)/(1+p^iq^jr^k))^{c(ij+jk+ki)}$$

where $c(i)$ is defined by $\sum_n c(n)q^n = \prod_{n>0}(1+q^n)/(1-q^n) = 1+2q+4q^2+8q^3+14q^4+\ldots$.

(7) Find all completely replicable functions. (The ones with integral coefficients were found by computer in [1].) Are nontrivial completely replicable functions always modular functions of genus 0? A proof of this conjecture of Norton's [28] which was not a case by case verification would be much neater than the argument in section 9.

(8) Is it possible to say anything interesting from Lie algebras constructed from the vertex algebras of lattices (other than the Leech lattice) as in section 6? The two obvious

candidates are the lattices E_8 and $E_8 \oplus E_8$, so that the corresponding Lie algebras have root lattices $II_{9,1}$ and $II_{17,1}$. The real simple roots of these Lie algebras are the Dynkin diagrams of the reflection groups of the lattices, which were described by Vinberg. There are some calculations connected with the Lie algebra of $II_{9,1}$ in [24].

References

[1]. D. Alexander, C. Cummins, J. McKay, C. Simons, Completely replicable functions, preprint.

[2]. M. F. Atiyah, K-theory, Benjamin, Inc. 1967.

[3]. R. E. Borcherds, Vertex algebras, Kac-Moody algebras, and the monster. Proc. Natl. Acad. Sci. USA. Vol. 83 (1986) 3068-3071.

[4]. R. E. Borcherds, Generalized Kac-Moody algebras. J. Algebra 115 (1988), 501–512.

[5]. R. E. Borcherds, Central extensions of generalized Kac-Moody algebras. J. Alg. 140, 330-335 (1991).

[6]. R. E. Borcherds, Lattices like the Leech lattice, J. Algebra, vol 130, No. 1, April 1990, 219-234.

[7]. R. E. Borcherds, J. H. Conway, L. Queen, N. J. A. Sloane, A monster Lie algebra?, Adv. Math. 53 (1984) 75-79. This paper is reprinted as chapter 30 of [12].

[8]. R. E. Borcherds, The monster Lie algebra, Adv. Math. Vol. 83, No. 1, Sept. 1990.

[9]. R. E. Borcherds, Vertex algebras, to appear.

[10]. H. Cartan, S. Eilenberg, Homological algebra, Princeton University Press 1956.

[11]. J. H. Conway, The automorphism group of the 26 dimensional even Lorentzian lattice. J. Algebra 80 (1983) 159-163. This paper is reprinted as chapter 27 of [12].

[12]. J. H. Conway, N. J. A. Sloane. Sphere packings, lattices and groups. Springer-Verlag New York 1988, Grundlehren der mathematischen Wissenschaften 290.

[13]. J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[14]. J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, R. A. Wilson, Atlas of finite groups, Clarendon Press, Oxford, 1985.

[15]. I. B. Frenkel, Representations of Kac-Moody algebras and dual resonance models, Applications of group theory in theoretical physics, Lect. Appl. Math. 21, A.M.S. (1985), p.325-353.

[16]. I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[17]. I. B. Frenkel, J. Lepowsky, A. Meurman, A natural representation of the Fischer-Griess monster with the modular function J as character, Proc. Natl. Acad. Sci. USA 81 (1984), 3256-3260.

[18]. I. B. Frenkel, Y-Z. Huang, J. Lepowsky, On axiomatic formulations of vertex operator algebras and modules, preprint.

[19]. I. B. Frenkel, H. Garland, G. Zuckerman, Semi-infinite cohomology and string theory, Proc. Natl. Acad. Sci. U.S.A. 83 (1986), 8442-8446.

[20]. H. Garland and J. Lepowsky, Lie algebra homology and the Macdonald-Kac formulas, Invent. math. 34 (1976), p.37-76.

[21]. P. Goddard and C. B. Thorn, Compatibility of the dual Pomeron with unitarity and the absence of ghosts in the dual resonance model, Phys. Lett., B 40, No. 2 (1972), 235-238.

[22]. R. C. Gunning, Lectures on modular forms, Annals of mathematical studies, Princeton University Press, 1962.

[23]. V. G. Kac, "Infinite dimensional Lie algebras", third edition, Cambridge University Press, 1990. (The first and second editions (Birkhauser, Basel, 1983, and C.U.P., 1985) do not contain the material on generalized Kac-Moody algebras that we need.)

[24]. V. G. Kac, R. V. Moody, M. Wakimoto, On E_{10}, preprint.

[25]. M. Koike, On Replication Formula and Hecke Operators, Nagoya University preprint.

[26]. B. Kostant, Lie algebra cohomology and the generalized Borel-Weil theorem. Annals of Math. 74, 329-387 (1961).

[27]. K. Mahler, On a class of non-linear functional equations connected with modular functions, J. Austral. Math. Soc. 22A, (1976), 65-118.

[28]. S. P. Norton, More on moonshine, Computational group theory, Academic press, 1984, 185-193.

[29]. S. P. Norton, Generalized Moonshine, Proc. Symp. Pure Math. 47 (1987) p. 208-209.

[30]. J. P. Serre, A course in arithmetic. Graduate texts in mathematics 7, Springer-Verlag, 1973.

[31]. J. G. Thompson, A finiteness theorem for subgroups of $PSL(2, \mathbf{R})$ which are commensurable with $PSL(2, \mathbf{Z})$. Proc. Symp. Pure Math. 37 (1979) p. 533-555.

arXiv:math/9801101v1 [math.QA] 21 Jan 1998

Modular Moonshine III.

11 October 1994, corrected 18 Nov 1997

Richard E. Borcherds,*
D.P.M.M.S., 16 Mill Lane, Cambridge, CB2 2SB, England.

Mathematics department, University of California at Berkeley, CA 94720-3840, U. S. A.

reb@dpmms.cam.ac.uk
www.dpmms.cam.ac.uk/~reb

In this paper we complete the proof of Ryba's modular moonshine conjectures [R] that was started in [B-R]. We do this by applying Hodge theory to the cohomology of the monster Lie algebra over the ring of p-adic integers in order to calculate the Tate cohomology groups of elements of the monster acting on the monster vertex algebra.

Contents.

1. Introduction.
Notation.
2. Representations of $\mathbf{Z}_p[G]$.
3. The $\mathbf{Z}[1/2]$-form of the monster Lie algebra.
4. The modular moonshine conjectures for $p \geq 13$.
5. Calculation of some cohomology groups.
6. Elements of type $3B$, $5B$, $7B$, and $13B$.
7. Open problems and conjectures.

1. Introduction.

This paper is a continuation of the earlier paper [B-R] so we will only briefly recall results that are discussed there in more detail. In [B-R] we constructed a modular superalgebra for each element of prime order in the monster, and worked out the structure of this superalgebra for some elements in the monster. In this paper we work out the structure of this superalgebra for the remaining elements of prime order.

Ryba conjectured [R] that for each element of the monster of prime order p of type pA there is a vertex algebra ${}^gV$ defined over the finite field $\mathbf{F}_p$ and acted on by the centralizer $C_M(g)$ of g in the monster group M, with the property that the graded Brauer character $\mathrm{Tr}(h|{}^gV) = \sum_n \mathrm{Tr}(h|{}^gV_n)q^n$ is equal to the Hauptmodul $\mathrm{Tr}(gh|V) = \sum_n \mathrm{Tr}(g|V_n)q^n$ (where V is the graded vertex algebra acted on by the monster constructed by Frenkel, Lepowsky and Meurman [FLM]). In [B-R] the vertex superalgebra ${}^gV$ was defined for any element $g \in M$ of odd prime order to be the sum of the Tate cohomology groups $\hat{H}^0(g, V[1/2]) \oplus \hat{H}^1(g, V[1/2])$ for a suitable $\mathbf{Z}[1/2]$ form $V[1/2]$ of V, and it was shown that $\mathrm{Tr}(h|{}^gV) = \mathrm{Tr}(h|\hat{H}^0(g, V[1/2])) - \mathrm{Tr}(h|\hat{H}^1(g, V[1/2]))$ was equal to the Hauptmodul of $gh \in M$. Hence to prove the modular moonshine conjecture for an element g of type pA it is enough to prove that $\hat{H}^1(g, V[1/2]) = 0$. In [B-R] this was shown by explicit calculation for the elements of type pA for $p \leq 11$, using the fact that these elements commute with an element of type $2B$. For $p \geq 13$ this method does not work as these elements do not commute with an element of type $2B$. The first main theorem of this paper is theorem 4.1 which states that if g is an element of prime order $p \geq 13$ not of type $13B$ then $\hat{H}^1(g, V[1/2]) = 0$ (assuming a certain condition about the monster Lie algebra, whose proof should appear later). Hence Ryba's conjectures are proved for all elements of M of type pA. (Actually this is not quite correct, because the proof for $p = 2$ in [B-R] assumes an unproved technical hypothesis.) The proof we give fails for exactly the cases already proved in [B-R].

We can also ask what happens for the other elements of order less than 13. The cases of elements of types $2B$ or $3C$ are also treated in [B-R]: for type $3C$, $\hat{H}^1(g, V[1/2])$ is again zero, and for type $2B$, $\hat{H}^1(g, V[1/2])$ is zero in even degree and $\hat{H}^0(g, V[1/2])$ is zero in odd degree. This leaves the cases where g is of type $3B$, $5B$, $7B$, or $13B$, when the Tate cohomology groups $\hat{H}^0(g, V[1/2])$ and $\hat{H}^1(g, V[1/2])$ are usually both nonzero in each degree. In these cases the structure of the cohomology groups is determined by our second main result, theorem 6.1, which states that the unique element σ of order 2 in the center of

---

* Supported by NSF grant DMS-9401186 and by a Royal Society research professorship.

$C_M(g)/O_p(C_M(g))$ acts as 1 on $\hat{H}^0(g,V[1/2])$ and as -1 on $\hat{H}^1(g,V[1/2])$. This determines the modular characters of both cohomology groups because $\mathrm{Tr}(h|\hat{H}^0(g,V[1/2]) + \mathrm{Tr}(h|\hat{H}^1(g,V[1/2]))$ is then given by $\mathrm{Tr}(h\sigma|\hat{H}^0(g,V[1/2])) - \mathrm{Tr}(h\sigma|\hat{H}^1(g,V[1/2]))$ which is the Hauptmodul of the element $gh\sigma \in M$. Hence both $\mathrm{Tr}(h|\hat{H}^0(g,V[1/2]))$ and $\mathrm{Tr}(h|\hat{H}^1(g,V[1/2]))$ can be written explicitly as linear combinations of two Hauptmoduls for any p-regular element $h \in C_M(g)$. We get two new modular vertex superalgebras acted on by double covers of sporadic groups when g has type $3B$ or $5B$: a vertex superalgebra over $\mathbf{F}_3$ acted on by $2.Suz$, and a vertex superalgebra over $\mathbf{F}_5$ acted on by $2.HJ$. For elements of types $7B$ and $13B$ the vertex superalgebras we get are acted on by the double covers $2.A_7$ and $2.A_4$ of alternating groups.

In particular if the integral form V of the monster vertex algebra discussed in [B-R] exists and has the properties conjectured there then all the cohomology groups $\hat{H}^i(g,V)$ are now known for all elements g of prime order in M.

We now discuss the proofs of the two main theorems. We first quickly dispose of the second main theorem: if an element $g \in M$ of odd order commutes with an element of type $2B$ then it is easy to work out its action on the $\mathbf{Z}[1/2]$-form of the monster vertex algebra and hence the Tate cohomology groups $\hat{H}^*(g,V[1/2])$ can be calculated by brute force, which is what we do in sections 5 and 6 for the elements of types $3B$, $5B$, $7B$, and $13B$.

The elements g of prime order $p \geq 13$ do not commute with any elements of type $2B$ (except when g has type $13B$ or $23AB$) so we cannot use the method above. Instead we adapt the proof of the moonshine conjectures in [B]. By using a $\mathbf{Z}_p$-form of the monster Lie algebra rather than a $\mathbf{Q}$-form we can find some complicated relations between the coefficients of the series $\sum_n \mathrm{Tr}(h|\hat{H}^0(g,V[1/2]_n))q^n$ and $\sum_n \mathrm{Tr}(h|\hat{H}^1(g,V[1/2]_n))q^n$. The reason we get more information by using $\mathbf{Z}_p$-forms rather than $\mathbf{Q}_p$-forms is that the group ring $\mathbf{Z}_p[\mathbf{Z}/p\mathbf{Z}]$ has 3 indecomposable modules which are free over $\mathbf{Z}_p$, while the group ring $\mathbf{Q}_p[\mathbf{Z}/p\mathbf{Z}]$ only has 2 indecomposable modules which are free over $\mathbf{Q}_p$, and this extra indecomposable module provides the extra information. The relations between the coefficients we get are similar to the relations defining completely replicable functions, except that some of the relations defining completely replicable functions (about "$(2p-1)/p^2$" of them) are missing. If p is large (≥ 13) we show that these equations have only a finite number of solutions. We then use a computer to find all the solutions, and see that the only solutions imply that $\hat{H}^1(g,V[1/2]) = 0$ for all g of prime order at least 13 other than those of type $13B$. This proof could probably be made to work for some smaller values of p such as 7 and 11, but the difficulty of showing that there are no extra solutions for the equations of the coefficients increases rapidly as p decreases, because for smaller p there are more equations missing.

Y. Martin [Ma] has recently found a conceptual proof that any completely replicable function is a modular function. Cummins and Gannon [CG] have greatly generalized this result using different methods, and showed that the functions are Hauptmoduls. It seems possible that their methods could be extended to replace the computer calculations in section 4, although Cummins has told me that this would probably require adding in some "missing" relations corresponding to the case $(p,mn) \neq 0$ in proposition 3.4.

We summarize the results of this paper and of [B-R] about the vertex superalgebras ${}^gV$ for all elements g of prime order p in the monster. If g is of type pA, or pB for $p > 13$, or $3C$, then ${}^gV$ is just a vertex algebra and not a superalgebra. If $p \geq 13$ this is proved in this paper by an argument which works for any self dual form of the monster vertex algebra but which relies on computer calculations and on a so far unpublished argument about integral forms of the monster Lie algebra. For $p < 13$ (or $p = 23$) this is proved in [B-R] by calculating explicitly for a certain $\mathbf{Z}[1/2]$ form of the monster vertex algebra; this avoids computer calculations but only works for one particular form of the monster vertex algebra. Also, if $p = 2$ the calculation depends on an as yet unproved technical assumption about the Dong-Mason-Montague construction of the monster vertex algebra from an element of type $3B$ ([DM] or [M]). If g is of type pB for $2 \leq p \leq 13$ then the super part of ${}^gV$ does not vanish. If $p = 2$ then ${}^gV$ is calculated explicitly in [B-R] (again using the unproved technical assumption) and it turns out that its ordinary part vanishes in odd degrees and its super part vanishes in even degrees. If $3 \leq p \leq 13$ then ${}^gV$ is calculated in this paper, and we find that there is an element in $C_M(g)$ acting as -1 on the super part and as 1 on the ordinary part of ${}^gV$. In all cases we have explicitly described the modular characters of $C_M(g)$ acting on the ordinary or super part of any degree piece of ${}^gV$ in terms of the coefficients of certain Hauptmoduls.

Notation.

$[A]$ is the element of K represented by the module A.

A, B, C	*G-modules.*
A_n	*The alternating group on n symbols.*
Aut	*The automorphism group of something.*
$c_g^+(n)$	*The n'th coefficient of the Hauptmodul of $g \in M$, equal to* $\mathrm{Tr}(1\vert^g V_n) = \dim(\hat{H}^0(g, V[1/2]_n)) - \dim(\hat{H}^1(g, V[1/2]_n))$.
$c_g^-(n)$	$\dim(\hat{H}^0(g, V[1/2]_n)) + \dim(\hat{H}^1(g, V[1/2]_n))$.
$c_{m,n}$	*Defined in proposition 4.3.*
$\mathbf{C}$	*The complex numbers.*
$C_M(g)$	*The centralizer of g in the group M.*
Co_1	*Conway's largest sporadic simple group.*
E	*The positive subalgebra of the monster Lie algebra.*
F	*The negative subalgebra of the monster Lie algebra.*
$\mathbf{F}_p$	*The finite field with p elements.*
f	*A homomorphism from K to $\mathbf{Q}$ defined in lemma 2.3.*
g	*An element of G, usually of order p.*
g_i	*The element g^i of G, used when it is necessary to distinguish the multiplication in the group ring from some other multiplication.*
$\langle g \rangle$	*The group generated by g.*
G	*A group, often cyclic of prime order p and generated by g.*
$h(A), h_n$	*See section 5.*
H	*The Cartan subalgebra of the monster Lie algebra.*
$\hat{H}^i(G, A)$	*A Tate cohomology group of the finite group G with coefficients in the G-module A.*
$\hat{H}^i(g, A)$	*means $\hat{H}^i(\langle g \rangle, A)$, where $\langle g \rangle$ is the cyclic group generated by g.*
$\hat{H}^*(g, A)$	*The sum of the Tate cohomology groups $\hat{H}^0(g, A)$ and $\hat{H}^1(g, A)$, considered as a super module.*
HJ	*The Hall-Janko sporadic simple group (sometimes denoted J_2).*
I	*The indecomposable module of dimension $p-1$ over $\mathbf{Z}_p[G]$, isomorphic to the kernel of the natural map from $\mathbf{Z}_p[G]$ to $\mathbf{Z}_p$ and to the quotient $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$.*
Im	*The image of a map.*
K	*A ring which is a free $\mathbf{Q}$-module with a basis of 3 elements $[\mathbf{Z}_p] = 1$, $[\mathbf{Z}_p[G]]$, and $[I]$.*
Ker	*The kernel of a map.*
$\Lambda, \hat{\Lambda}$	*The Leech lattice and a double cover of the Leech lattice.*
$\Lambda^n(A)$	*The n'th exterior power of A.*
$\Lambda^*(A)$	*The exterior algebra $\oplus_n \Lambda^n(A)$ of A.*
L	*An even lattice.*
$\mathbf{m}$	*The monster Lie algebra.*
M	*The monster simple group.*
N_G	*The element $\sum_{g \in G} g$ of $\mathbf{Z}_p[G]$.*
$O_p(G)$	*The largest normal p-subgroup of the finite group G.*
p	*A prime, usually the order of g.*
$\mathbf{Q}, \mathbf{Q}_p$	*The rational numbers and the field of p-adic numbers.*
ρ	*The Weyl vector of the monster Lie algebra.*
$\mathbf{R}$	*The real numbers.*
R_p	*A finite extension of the p-adic integers.*
$R(i)$	*Defined just before lemma 3.2.*
σ	*An automorphism of type $2B$ in the monster or the automorphism -1 of the Leech lattice.*
S_n	*A symmetric group.*
$S^n(A)$	*The n'th symmetric power of A.*
$S^*(A)$	*The symmetric algebra $\oplus_n S^n(A)$ of A.*
Suz	*Suzuki's sporadic simple group.*
Tr	$\mathrm{Tr}(g\vert A)$ *is the usual trace of g on a module A if A is a module over a ring of characteristic 0, and the Brauer trace if A is a module over a field of finite characteristic.*
$V[1/n]$	*A $\mathbf{Z}[1/n]$-form of the monster vertex algebra.*
V_Λ	*The integral form of the vertex algebra of $\hat{\Lambda}$.*

V_n *The degree n piece of V.*
V^n *An eigenspace of some group acting on V.*
${}^gV$ *A modular vertex algebra or superalgebra given by $\hat{H}^*(g, V[1/n])$ for some n coprime to $|g|$.*
$\mathbf{Z}$ *The integers.*
$\mathbf{Z}_p$ *The ring of p-adic integers.*
ω *A cube root of 1 or a conformal vector.*
Ω *The Laplace operator on $\Lambda^*(E)$.*

2. Representations of $\mathbf{Z}_p[G]$.

We give some auxiliary results about modules over $\mathbf{Z}_p[G]$ where G is a cyclic group generated by g of prime order p, and in particular calculate the exterior and symmetric algebras of all indecomposable modules. All modules will be free over $\mathbf{Z}_p$ and will be either finitely generated or $\mathbf{Z}$ graded with finitely generated pieces of each degree. All tensor products will be taken over $\mathbf{Z}_p$.

Recall from [B-R section 2] that there are 3 indecomposable modules over $\mathbf{Z}_p[G]$, which are $\mathbf{Z}_p$, the group ring $\mathbf{Z}_p[G]$, and the module I of $\mathbf{Z}_p[G]$ that is isomorphic to the kernel of the natural map from $\mathbf{Z}_p[G]$ to $\mathbf{Z}_p$ and to the quotient $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$ (where $N_G = 1 + g + g^2 + \cdots + g^{p-1}$). Their Tate cohomology groups are given by $\hat{H}^0(g, \mathbf{Z}_p) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g, \mathbf{Z}_p) = 0$, $\hat{H}^0(g, \mathbf{Z}_p[G]) = 0$, $\hat{H}^1(g, \mathbf{Z}_p[G]) = 0$, $\hat{H}^0(g, I) = 0$, and $\hat{H}^1(g, I) = \mathbf{Z}/p\mathbf{Z}$.

Lemma 2.1. *The tensor products of these modules are given as follows:*

$$\begin{aligned}\mathbf{Z}_p \otimes X &= X \ \textit{(for any } X\textit{)}\\ \mathbf{Z}_p[G] \otimes X &= \mathbf{Z}_p[G]^{\dim(X)} \ \textit{(the sum of } \dim(X) \textit{ copies of } \mathbf{Z}_p[G]\textit{)}\\ I \otimes I &= \mathbf{Z}_p[G]^{p-2} \oplus \mathbf{Z}_p.\end{aligned}$$

Proof. The case $\mathbf{Z}_p \otimes X$ is trivial. If X has a basis $x_1, \ldots, x_n$ then for any fixed i the elements $g^j \otimes g^j(x_i)$ ($0 \le j < p$) form a basis for a submodule of $\mathbf{Z}_p[G] \otimes X$ isomorphic to $\mathbf{Z}_p[G]$, and $\mathbf{Z}_p[G] \otimes X$ is the direct sum of these submodules for all i, which proves the result about $\mathbf{Z}_p[G] \otimes X$. For the case $I \otimes I$ we look at the cohomology sequence of the short exact sequence

$$0 \to I \otimes I \to I \otimes \mathbf{Z}_p[G] \to I \otimes \mathbf{Z}_p \to 0$$

to see that $\hat{H}^0(I \otimes I) = \mathbf{Z}/p\mathbf{Z}$ and $\hat{H}^1(I \otimes I) = 0$, which implies that $I \otimes I$ must be the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$. The number of copies can be worked out by looking at the dimensions of both sides. This proves lemma 2.1.

Corollary 2.2. *If A and B are $\mathbf{Z}_p[G]$-modules then $\hat{H}^*(g, A \otimes B) = \hat{H}^*(g, A) \otimes \hat{H}^*(g, B)$.*

Proof. Check each of the 9 cases when A and B are indecomposable modules using lemma 2.1. This proves corollary 2.2.

We form the ring K which is a free $\mathbf{Q}$-module with a basis of 3 elements $[\mathbf{Z}_p] = 1$, $[\mathbf{Z}_p[G]]$, and $[I]$ corresponding to the indecomposable modules of $\mathbf{Z}_p[G]$, whose addition and multiplication are those induced by tensor products and direct sums of modules. If A is a $\mathbf{Z}_p[G]$-module then we write $[A]$ for the element of K represented by A. Notice that this is not the same as the (tensor product with $\mathbf{Q}$ of the) Grothendieck ring of finitely generated projective $\mathbf{Z}_p[G]$ modules because we do not assume the relations given by nonsplit short exact sequences. The Grothendieck ring (tensored with $\mathbf{Q}$) is a quotient of our ring K by the ideal generated by $[\mathbf{Z}_p] + [I] - [\mathbf{Z}_p[G]]$.

Lemma 2.3. *There are exactly 3 ring homomorphisms* $\dim$, $\mathrm{Tr}(g|.)$, *and f from K to $\mathbf{Q}$ (which correspond to the 3 elements of the spectrum of K). They are given by*

1. $\dim([\mathbf{Z}_p]) = 1$, $\dim([\mathbf{Z}_p[G]]) = p$, $\dim([I]) = p - 1$.
2. $\mathrm{Tr}(g|[\mathbf{Z}_p]) = 1$, $\mathrm{Tr}(g|[\mathbf{Z}_p[G]]) = 0$, $\mathrm{Tr}(g|[I]) = -1$.
3. $f([\mathbf{Z}_p]) = 1$, $f([\mathbf{Z}_p[G]]) = 0$, $f([I]) = 1$.

Proof. From corollary 2.2 we know that the products in the ring K are given by $[\mathbf{Z}_p[G]][\mathbf{Z}_p[G]] = p[\mathbf{Z}_p[G]]$, $[\mathbf{Z}_p[G]][I] = (p-1)[\mathbf{Z}_p[G]]$, and $[I][I] = (p-2)[\mathbf{Z}_p[G]] + 1$. Hence if $x \in \mathbf{Q}$ and $y \in \mathbf{Q}$ are the images of $[\mathbf{Z}_p[G]]$ and I under some homomorphism they must satisfy $x^2 = px$, $xy = (p-1)x$, and $y^2 = (p-2)x+1$. The only solutions to these equations are $(x,y) = (p, p-1)$, $(0,1)$, or $(0,-1)$. This proves lemma 2.3.

The first two homomorphisms in lemma 2.3 are the homomorphisms which factor through the Grothendieck ring and will not be very interesting to us. We can use them in place of f in the arguments in the rest of the paper, but this only leads to relations between the coefficients of $\sum_n \dim(V_n)q^n$ and $\sum_n \mathrm{Tr}(g|V_n)q^n$, which are essentially the relations already used in [B] to determine these functions. It is the third homomorphism f which detects the integral structure on $V[1/2]$ which will give us enough extra information to prove theorem 4.1.

In general, taking the trace of any fixed element of a group gives a homomorphism from a representation ring of a group to some field, and quite often (e.g., for the complex representation ring of a group) this gives a 1:1 correspondence between homomorphisms and (certain sorts of) conjugacy classes of the group. Lemma 2.3 gives an example of a homomorphism f not constructed in this way.

A permutation module is a module for G with a basis acted on by G. We recall from [B-R] that a module A is a permutation module if and only if $\hat{H}^1(g, A) = 0$, which in turn is equivalent to saying that A is a sum of copies of $\mathbf{Z}_p$ and $\mathbf{Z}_p[G]$.

Lemma 2.4. *If A is a permutation module over any ring for any finite group g then so are the symmetric powers $S^n(A)$. If all elements of G act on a basis of A as products of cycles of odd length then the exterior powers $\Lambda^n(A)$ are also permutation modules.*

Proof. For symmetric powers this is proved in [B-R]. For the proof for exterior powers, let $a_1, \ldots, a_n$ be a basis for A acted on by G. Then G acts on the set of all elements of the form $\pm a_{i_1} \wedge \cdots \wedge a_{i_n}$, which consists of all elements of a basis together with their negatives. If the action of G on the basis of A has the property that all elements of G are products of cycles of odd length, then there is no element of G that maps any element of the form $\pm a_{i_1} \wedge \cdots \wedge a_{i_n}$ to its negative. This means that we can find a G-invariant subset of these elements which forms a basis for $\Lambda^n(A)$. This proves lemma 2.4.

Lemma 2.5. *The exterior powers of indecomposable modules are given by*

$$\begin{aligned}
\Lambda^0(\mathbf{Z}_p) &= \mathbf{Z}_p \\
\Lambda^1(\mathbf{Z}_p) &= \mathbf{Z}_p \\
\Lambda^n(\mathbf{Z}_p) &= 0 \textit{ if } n \neq 0, 1 \\
\Lambda^0(\mathbf{Z}_p[G]) &= \mathbf{Z}_p \\
\Lambda^p(\mathbf{Z}_p[G]) &= \mathbf{Z}_p \textit{ if } p \textit{ is odd, and } I \textit{ if } p = 2 \\
\Lambda^n(\mathbf{Z}_p[G]) &= \mathbf{Z}_p[G]^{\binom{p-1}{n}} \textit{ if } 1 \leq n \leq p-1 \\
\Lambda^n(\mathbf{Z}_p[G]) &= 0 \textit{ if } n > p \\
\Lambda^n(I) &= \mathbf{Z}_p \oplus \mathbf{Z}_p[G]^{(\binom{p-1}{n}-1)/p} \textit{ if } n \textit{ is even and } 0 \leq n < p \\
\Lambda^n(I) &= I \oplus \mathbf{Z}_p[G]^{(\binom{p-1}{n}+1-p)/p} \textit{ if } n \textit{ is odd and } 0 \leq n < p \\
\Lambda^n(I) &= 0 \textit{ if } n \geq p
\end{aligned}$$

Proof. For the exterior powers of $\mathbf{Z}_p[G]$ with p odd we apply lemma 2.4 to see that all exterior powers are sums of $\mathbf{Z}_p$ and $\mathbf{Z}_p[G]$, and the exterior powers can then be identified by checking the dimension and the trace of g. The cohomology groups of the exterior powers of I can then be worked out by induction using the exact sequence

$$0 \to \Lambda^n(I) \to \Lambda^n(\mathbf{Z}_p[G]) \to \Lambda^{n-1}(I) \to 0.$$

The cohomology groups determine how many copies of $\mathbf{Z}_p$ and I occur in the decomposition into indecomposable modules, so we can work out the exterior powers of I from this. The remaining cases of lemma 2.5 are trivial, so this proves lemma 2.5.

These exterior powers induce operators on the ring K which we also denote by Λ^n. For example, we can use these operators to define Adams operations ψ^n on the ring K by

$$\sum_{n \in \mathbf{Z}} (-1)^n \Lambda^n([A])q^n = \exp(-\sum_{n \in \mathbf{Z}} \psi^n([A])q^n/n).$$

Corollary 2.6. *If p is odd then*

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(\mathbf{Z}_p)])q^n = 1-q$$

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(\mathbf{Z}_p[G])])q^n = 1-q^p$$

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(I)])q^n = (1+q^p)/(1+q)$$

Proof. This follows immediately from lemma 2.5 and lemma 2.3.

Lemma 2.7. *The symmetric powers of indecomposable modules are given by*

$$S^n(\mathbf{Z}_p) = \mathbf{Z}_p \text{ if } n\geq 0$$

$$S^n(\mathbf{Z}_p[G]) = \mathbf{Z}_p[G]^{\binom{p+n-1}{n}/p} \text{ if } (p,n)=1$$

$$S^n(\mathbf{Z}_p[G]) = \mathbf{Z}_p[G]^{(\binom{p+n-1}{n}-1)/p}\oplus\mathbf{Z}_p \text{ if } p|n$$

$$S^n(I) = \mathbf{Z}_p\oplus\mathbf{Z}_p[G]^{(\binom{p+n-2}{n}-1)/p} \text{ if } p|n$$

$$S^n(I) = I\oplus\mathbf{Z}_p[G]^{(\binom{p+n-2}{n}+1-p)/p} \text{ if } n\equiv 1 \bmod p$$

$$S^n(I) = \mathbf{Z}_p[G]^{(\binom{p+n-2}{n})/p} \text{ if } n\not\equiv 0,1 \bmod p$$

Proof. For symmetric powers of $\mathbf{Z}_p$ this is trivial and for symmetric powers of $\mathbf{Z}_p[G]$ it can be proved in the same way as in lemma 2.5. If $p=2$ then the case of I is easy to do as I is just one dimensional. For the case of I when p is odd we let N_G be the element $1+g+\cdots+g^{p-1}\in\mathbf{Z}_p[G]$ and consider the quotient of the symmetric algebra $S^*(\mathbf{Z}_p[G])$ by the ideal (N_G). This is the universal commutative algebra generated by $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$, so it is $S^*(\mathbf{Z}_p[G]/N_G\mathbf{Z}_p)$ which is isomorphic to $S^*(I)$ because $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$ is isomorphic to I. Therefore there is an exact sequence

$$0\rightarrow S^{n-1}(\mathbf{Z}_p[G])\rightarrow S^n(\mathbf{Z}_p[G])\rightarrow S^n(I)\rightarrow 0.$$

Using the known cohomology groups of the symmetric powers of $\mathbf{Z}_p[G]$ and the exact sequence of cohomology groups of this exact sequence gives the cohomology groups of $S^n(I)$ and hence the number of times that $\mathbf{Z}_p$ and I occur in the decomposition of $S^n(I)$. This proves lemma 2.7.

Corollary 2.8. *The cohomology rings of the symmetric algebras of indecomposable modules are give as follows.*

$\hat{H}^*(S^*(\mathbf{Z}_p))$ *is a polynomial algebra with one ordinary generator of degree 1.*

$\hat{H}^*(S^*(\mathbf{Z}_p[G]))$ *is a polynomial algebra with one ordinary generator of degree p.*

$\hat{H}^*(S^*(I))$ *is the tensor product of a polynomial algebra with one ordinary generator of degree p and an exterior algebra with one super generator of degree 1 if p is odd. (If p is even then* $\hat{H}^*(S^*(I))$ *is a polynomial algebra on one super generator of degree 1, but we do not need this case.)*

Proof. This follows from the calculation of the symmetric powers in lemma 2.7. The product structure on the symmetric algebras can be worked out using the formula $d(ab)=(da)b\pm a(db)$ relating the cup product with the map d from $\hat{H}^i(C)$ to $\hat{H}^{i+1}(A)$ associated to the exact sequence $0\rightarrow A\rightarrow B\rightarrow C\rightarrow 0$. This proves corollary 2.8.

Lemma 2.9. *Suppose that* $A_{-1}=0, A_0,\ldots,A_n,A_{n+1}=0$ *is a sequence of modules and we have maps* d *from* A_i *to* A_{i+1} *and* d^* *from* A_{i+1} *to* A_i *for all* i *such that* $d^2=0$, $d^{*2}=0$, *and* $dd^*+d^*d=k$ *for some fixed integer* k *coprime to* p. *Then*

$$[A_0]-[A_1]+[A_2]-\cdots+(-1)^n[A_n]=0.$$

Proof. The fact that k is a unit in $\mathbf{Z}_p$ implies that A_i is the direct sum $dd^*(A_i)\oplus d^*d(A_i)$ and that d is an isomorphism from $d^*d(A_i)$ to $dd^*(A_{i+1})$. The lemma follows from this because the sequence splits as the sum of these isomorphisms. (Notice that the relation between the $[A_i]$'s does not follow just from the exactness of d; for example, look at the short exact sequence $0\rightarrow I\rightarrow\mathbf{Z}_p[G]\rightarrow\mathbf{Z}_p\rightarrow 0$. It is easy to check that in this case there is a map d^* with $d^{*2}=0$ and $dd^*+d^*d=p$, so the condition that $(k,p)=1$ is also necessary.) This proves lemma 2.9.

Lemma 2.10. *Suppose that A is a finitely generated free module over $\mathbf{Z}$ or $\mathbf{Z}_p$ with a symmetric self dual form (,). Then this induces a symmetric self dual form (defined in the proof) on any exterior power $\Lambda^n(A)$.*

Proof. We define the bilinear form on $\Lambda^n(A)$ by

$$(a_1 \wedge \cdots \wedge a_n, b_1 \wedge \cdots \wedge b_n) = \sum_{\sigma \in S_n} \chi(\sigma)(a_1, b_{\sigma(1)}) \cdots (a_n, b_{\sigma(n)})$$

where $\chi(\sigma)$ is 1 or -1 depending on whether σ is an even or odd permutation. This obviously defines a symmetric bilinear form on $\Lambda^n(A)$ with values in $\mathbf{Z}$ or $\mathbf{Z}_p$, and we have to check it is unimodular. The determinant of $\Lambda^n(A)$ is equal to $a \det(A)^b$ for some constants a and b depending on n, but not on A. We can work out a by letting A be the lattice with a base of orthogonal elements $a_1, a_2, \ldots$ with norms ± 1, when we see that $\Lambda^n(A)$ also has an orthogonal base of elements $a_{i_1} \wedge a_{i_2} \cdots$ with $i_1 < i_2 < \cdots$ with norms ± 1, so both A and $\Lambda^n(A)$ have discriminant 1 and so $a = 1$. Hence if A is self dual then $\Lambda^n(A)$ has discriminant 1 and must also be self dual. This proves lemma 2.10. (It is easy to check that the constant b is equal to $\dim(\Lambda^n(A))/\dim(A)$ but we do not need this.)

The analogue of this lemma for symmetric powers is false. The proof breaks down because the element $a_{i_1} a_{i_2} \cdots$ with $i_1 \leq i_2 \leq \cdots$ does not have norm ± 1 if some of the i_j's are equal. In section 5 we will construct a graded algebra $h(A)$, such that the homogeneous components of $h(A)$ have self dual symmetric bilinear forms if A does, and $h(A) \otimes \mathbf{Q}_p$ is a sum of tensor products of symmetric powers of A.

The following corollary is purely for entertainment and is not used elsewhere in the paper.

Corollary 2.11. *There is a 28-dimensional self dual positive definite lattice with no roots, acted on by the orthogonal group $O_8^+(\mathbf{F}_2)$.*

Proof. By lemma 2.10 the lattice $\Lambda^2(E_8)$ is self dual, and is obviously acted on by the automorphism group of the E_8 lattice modulo -1, which is just $O_8^+(\mathbf{F}_2)$. It is easy to check that this lattice has no roots either by calculation or by observing that there is no suitable root lattice acted on by $O_8^+(\mathbf{F}_2)$. This proves lemma 2.11.

The lattice $\Lambda^2(E_8)$ has theta function

$$\sum_{\lambda \in \Lambda^2(E_8)} q^{\lambda^2/2} = 1 + 2240q^{3/2} + 98280q^{4/2} + 1790208q^{5/2} + 19138560q^{6/2} + \cdots.$$

The vectors of norm 3 are just the 2240 vectors generating one of the 1-dimensional sublattices of the form $\Lambda^2(L)$ where L is one of the 1120 sublattices of E_8 isomorphic to A_2.

3. The $\mathbf{Z}[1/2]$-form of the monster Lie algebra.

We find some relations between the dimensions of the Tate cohomology groups $\hat{H}^i(g, V[1/2])$ by studying the Lie algebra cohomology of the positive subalgebra of a $\mathbf{Z}_p$-form of the monster Lie algebra. In the next section we will use these relations to compute the cohomology groups in the case when g has order at least 13. We recall that $V[1/2]$ is the self dual $\mathbf{Z}[1/2]$-form of Frenkel, Lepowsky, and Meurman's monster vertex algebra [FLM] described in [B-R] (but the arguments in this section work for any self dual $\mathbf{Z}_p$-form of the monster vertex algebra).

This section should really be about the integral form of the monster Lie algebra, but this has not been constructed yet because of the lack of a construction of a self dual integral form of the monster vertex algebra. However, as we are only looking at the action of elements of odd order p in the monster and tensoring everything with $\mathbf{Z}_p$, a $\mathbf{Z}[1/2]$-form is just as good as an integral form.

We briefly recall some properties of the monster Lie algebra $\mathbf{m}$ from [B]. This is a $II_{1,1}$-graded Lie algebra acted on by the monster, where $II_{1,1}$ is the 2-dimensional even self dual Lorentzian lattice whose element (m, n) has norm $-2mn$. The piece of degree (m, n) of $\mathbf{m}$ is isomorphic to V_{mn} if $(m, n) \neq (0, 0)$ and to $\mathbf{R}^2$ if $(m, n) = (0, 0)$. The Lie algebra $\mathbf{m}$ is the sum of 3 subalgebras E, F, and H, where the Cartan subalgebra H is the piece of degree $(0, 0)$ of $\mathbf{m}$, E is the sum of all pieces of degree (m, n) of $\mathbf{m}$ for $m > 0$, and F is the sum of all pieces of degree (m, n) of $\mathbf{m}$ for $m < 0$.

The $\mathbf{Z}[1/2]$ form on V induces a self dual $\mathbf{Z}[1/2]$-form and hence a $\mathbf{Z}_p$-form on $\mathbf{m}$. From now on we will write $\mathbf{m}[1/2]$ for this $\mathbf{Z}[1/2]$-form. In the preprint version of this paper I implicitly assumed that the degree (m, n) piece of $\mathbf{m}$ was self dual and isomorphic as a module over the monster to V_{mn}. At the last moment I realized that this is not at all clear. The problem is that although the no-ghost theorem gives an isomorphism of both spaces tensored with $\mathbf{Q}$, and both spaces have $\mathbf{Z}[1/2]$-forms, there is no obvious reason why this isomorphism should map one $\mathbf{Z}[1/2]$ form to the other. (In fact for the fake monster Lie algebra the corresponding isomorphism does not preserve the integral forms.) Fortunately in this paper we only need the following weaker statement:

Assumption. *If $m < p$ then the degree (m, n) piece of $\mathbf{m} \otimes \mathbf{Z}_p$ is self dual under the natural bilinear form and isomorphic to $V_{mn} \otimes \mathbf{Z}_p$ as a $\mathbf{Z}_p$ module acted on by the monster.*

I believe I have a proof of this; if all goes well it will appear in a paper provisionally titled "The fake monster formal group". Until this paper appears all the statements in sections 3 and 4 and the modular moonshine conjectures for elements of type pA ($p \geq 13$) should have this assumption added as a hypothesis.

I do not know if the condition $m < p$ in the assumption above is necessary; if it is not, then this condition (usually added in parentheses) can be missed out of some of the following lemmas. It seems very likely that self duality still holds without it, but I am not so sure about the action of the monster being the same if both m and n are divisible by p.

The monster Lie algebra has a Weyl vector ρ, i.e., a vector in the root lattice $II_{1,1}$ such that $(\rho, \alpha) = -(\alpha, \alpha)/2$ for every simple root α. This follows from [B, theorem 7.2] which states that this simple roots are the vectors $(1, n)$ for $n \geq 0$ or $n = -1$, with multiplicities $c(n)$. Hence the Weyl vector is $\rho = (1, 0)$.

We recall some facts about Hodge theory applied to Lie algebra cohomology. The self dual symmetric bilinear form on the monster Lie algebra identifies E with the dual space of F, and using the Cartan involution which maps F to E we see that we have a symmetric bilinear form on the Lie algebra E (which is the restriction of the contravariant form of the monster Lie algebra to E) and in particular we can identify E with its dual. The Lie algebra cohomology is the cohomology of an operator d which maps each space $\Lambda^i(E)$ to $\Lambda^{i+1}(E)$. The adjoint of this operator is denoted by d^*, and we define the Laplace operator Ω by $\Omega = dd^* + d^*d$. As for any generalized Kac-Moody algebra with a Weyl vector, the action of the Laplace operator $\Omega = dd^* + d^*d$ on the degree $\alpha = (m, n) \in II_{1,1}$ piece of $\Lambda^*(E)$ is given by multiplication by $(\alpha, \alpha + 2\rho)/2 = (m-1)n$. The Lie algebra cohomology group $H^i(E)$ of E can be identified with the zero eigenspace of Ω on $\Lambda^i(E)$. The Laplace operator restricted to $\Lambda^1(E) = E$ is half the partial Casimir operator Ω_0 defined in [K, section 2.5]; the extra factor of $1/2$ is necessary to make things work well over rings not containing $1/2$, but this is not important in this paper as we work over the ring of p-adic integers for odd p which always contains $1/2$.

Proposition 3.1. *The operators d and d^* act on the $\mathbf{Z}[1/2]$-form $\Lambda^*(E[1/2]) \subset \Lambda^*(E)$ of the exterior algebra (in all degrees (m, n) with $m < p$).*

Proof. The main point is that all homogeneous pieces of the $\mathbf{Z}[1/2]$-form of the exterior algebra are self dual under the contravariant bilinear form. This follows from lemma 2.10 and the fact that each homogeneous component is a finite sum of tensor products of exterior powers of the homogeneous components of $E[1/2]$. The result now follows from the fact that if d is a homomorphism of finite dimensional free $\mathbf{Z}[1/2]$-modules from A to B and A and B have self dual symmetric inner products then the adjoint d^* from B to A is well defined. This proves proposition 3.1.

We let $R(0)$ be the ring of formal Laurent series $K[[r, q]][1/r, 1/q]$ with coefficients in K, and we let $R(1)$ be the space of power series in $R(0)$ only involving monomials $r^m q^n$ with $p|mn$, and we let $R(2)$ be the subring of $R(0)$ of power series only involving powers of r^p and q^p. So we have the inclusions $R(2) \subset R(1) \subset R(0)$ and $R(1)$ is a module over $R(2)$ (but is not a ring itself). (We can also define $R(0)$, $R(1)$, and $R(2)$ for arbitrary generalized Kac-Moody algebras, in particular for the fake monster Lie algebra: $R(0)$ is a completion of the group ring of the root lattice, $R(1)$ is a subspace corresponding to the elements α of the root lattice with $p|(\alpha, \alpha)/2$, and $R(2)$ is the subring corresponding to p times the root lattice.) If A is a $\mathbf{Z}[1/2]$-module then we write $[A]$ for $[A \otimes_{\mathbf{Z}[1/2]} \mathbf{Z}_p]$.

Lemma 3.2. *If $(p, (m-1)n) = 1$ (and $m < p$) then the piece of degree (m, n) of*

$$[\Lambda^*(E[1/2])] = [\Lambda^0(E[1/2])] - [\Lambda^1(E[1/2])] + [\Lambda^2(E[1/2])] - \cdots$$

vanishes. In other words

$$r^{-1}\Lambda^*(\sum_{m>0,n\in\mathbf{Z}} r^m q^n[V_{mn}])$$

lies in $R(1)$.

Proof. The Laplace operator $\Omega = dd^* + d^*d$ acts as multiplication by $(m-1)n$ on the degree (m,n) piece of $\Lambda^*(E[1/2])$, so the result follows from lemma 2.9 if we let A_i be the degree (m,n) piece of $\Lambda^i(E[1/2])$. This proves lemma 3.2.

Lemma 3.3. *If A is a $\mathbf{Z}_p[G]$-module with $\mathrm{Tr}(g|A) = c^+$, $f([A]) = c^-$, (and $m < p$) then*

$$f(\Lambda^*(r^m q^n A)) = (1 - r^m q^n)^{(c^+ + c^-)/2}(1 + r^m q^n)^{(c^+ - c^-)/2} U$$

where U is a unit of the ring $R(2)$.

Proof. It is sufficient to prove this for an indecomposable module A because both sides are multiplicative. We just check each of the three possibilities for A using lemma 2.6.

If $A = \mathbf{Z}_p$ then $c^+ = c^- = 1$ and $f(\Lambda^*(r^m q^n[\mathbf{Z}_p])) = 1 - r^m q^n$, so lemma 3.3 is true in this case.

If $A = \mathbf{Z}_p[G]$ then $c^+ = c^- = 0$ and $f(\Lambda^*(r^m q^n[\mathbf{Z}_p[G]])) = 1 - r^{mp} q^{np}$ which is a unit in $R(2)$, so lemma 3.3 is true in this case.

If $A = I$ then $c^+ = -1$, $c^- = 1$ and

$$f(\Lambda^*(r^m q^n[I])) = (1 + r^{mp} q^{np})/(1 + r^m q^n) = U(1 + r^m q^n)^{-1},$$

so lemma 3.3 is also true in this case. This proves lemma 3.3.

We define integers $c_g^+(n)$ and $c_g^-(n)$ by

$$\begin{aligned}
c_g^+(n) &= \mathrm{Tr}(g|V_n) = \dim(\hat{H}^0(g, V[1/2]_n)) - \dim(\hat{H}^1(g, V[1/2]_n))\\
c_g^-(n) &= f(V[1/2]_n) = \dim(\hat{H}^0(g, V[1/2]_n)) + \dim(\hat{H}^1(g, V[1/2]_n)).
\end{aligned}$$

We use the numbers $c_g^+(n)$ and $c_g^-(n)$ rather than the apparently simpler numbers $\dim(\hat{H}^0(g, V[1/2]_n))$ and $\dim(\hat{H}^1(g, V[1/2]_n))$ because they turn out to be coefficients of Hauptmoduls, and because $c_g^+(n)$ is already known.

Proposition 3.4. *If $(p, mn) = 1$ (and $m < p$) then the coefficient of $r^m q^n$ in*

$$r^{-1} \prod_{m>0,n\in\mathbf{Z}} (1 - r^m q^n)^{(c_g^-(mn) + c_g^+(mn))/2}(1 + r^m q^n)^{(c_g^+(mn) - c_g^-(mn))/2}$$

vanishes. In other words this power series lies in $R(1)$.

Proof. We apply the homomorphism f to the expression in lemma 3.2 and use lemma 3.3. We find that the expression in this proposition is equal to an element of $R(1)$ times a unit in $R(2)$, and is therefore still an element of $R(1)$. Therefore its coefficients of $r^m q^n$ vanish unless $p|mn$. This proves proposition 3.4.

Proposition 3.4 can be generalized in the obvious way to any generalized Kac-Moody algebra which has a self dual $\mathbf{Z}_p$-form, a Weyl vector, and an integral root lattice.

We can also ask whether or not the coefficients of $r^m q^n$ in proposition 3.4 vanish when $p|mn$. Some numerical calculations suggest that they usually do not.

4. The modular moonshine conjectures for $p \geq 13$.

In this section we prove the following theorem, which completes the proof of the modular moonshine conjectures of [R section 6] (apart from a small technicality in the case $p = 2$.)

Theorem 4.1. *If $g \in M$ is an element of prime order $p \geq 17$ or an element of type $13A$ then $\hat{H}^1(g, V[1/2]) = 0$.*

This implies that ${}^gV = \hat{H}^0(g, V[1/2]) = \hat{H}^*(g, V[1/2])$ is a vertex algebra whose modular character is given by Hauptmoduls, and whose homogeneous components have the characters of [R definition 2].

The proof of this theorem will occupy the rest of this section. We have to show that the numbers $c_g^-(n)$ of proposition 3.4 are equal to the numbers $c_g^+(n)$, because the difference is twice the dimension of $\hat{H}^1(g, V[1/2]_n)$. We start by summarizing what we know about these numbers.

Lemma 4.2.

1. *The numbers $c_g^-(n)$ and $c_g^+(n)$ satisfy the relations given in proposition 3.4.*
2. *The numbers $c_g^-(n)$ are integers, with $c_g^-(n) \equiv c_g^+(n) \bmod 2$.*
3. *$c_g^-(n) \geq |c_g^+(n)|$ and $(p-2)c_g^+(n) + pc_g^-(n) \leq 2c_1^+(n)$, (where $c_1^+(n)$ is the coefficient of q^n in the elliptic modular function).*
4. *The numbers $c_g^+(n)$ are the coefficients of the Hauptmodul of the element $g \in M$.*

Proof. These properties follow easily from proposition 3.4 together with the fact that $\mathbf{Z}_p \otimes V_n$ has dimension $c_1^+(n)$ and is the sum of $(c_g^-(n) + c_g^+(n))/2$ copies of $\mathbf{Z}_p$, $(c_g^-(n) - c_g^+(n))/2$ copies of I, and some copies of $\mathbf{Z}_p[G]$. This proves lemma 4.2.

We will prove theorem 4.1 by showing that if g satisfies the conditions of theorem 4.1 then the conditions 1 to 4 in lemma 4.2 imply that $c_g^-(n) = c_g^+(n)$. We find a finite set of possible solutions of conditions 1, 2, and 4 of lemma 4.2, and find that for each p only one of these satisfies the condition 3. (There are often several solutions not satisfying condition 3.) The proof is just a long messy calculation and the reader should not waste time looking at it.

We first put the relations into a more convenient form.

Proposition 4.3. *There are integers $c_{m,n}$ defined for $m, n > 0$ (and $m < p$) with the following properties.*

1.
$$-\log\left(1 + \sum_{m>0,n>0} c_{m,n} r^m q^n\right) \equiv \sum_{m>0,n>0} \sum_{d|(m,n)} \frac{c_g^{(-)^d}(mn/d^2)}{d} r^m q^n \bmod r^p$$

2. *If $(p, mn) = 1$ then $c_{m+1,n} = c_{m,n+1}$.*

Proof. This follows from proposition 3.4 if we multiply both sides by $(r-q)/rq$ and then take the logarithm of both sides. This proves proposition 4.3.

Lemma 4.4. *If $p > 13$ then the values of $c_g^-(n)$ for $n \leq 21$, $c_g^-(2n)$ for $2n \leq 32$, and $c_g^-(36)$ and $c_g^-(45)$ are given by polynomials in the numbers $c_g^-(i)$ for $i = 1, 2, 4, 5$ with coefficients in $\mathbf{Z}_3$.*

Proof. We can evaluate the elements $c(n), n \leq 21$ and $c(2n), n \leq 16$ using the argument for cases 1 and 2 of lemma 4.7 below. We can evaluate $c(36)$ by looking at the coefficient of r^4q^9 in proposition 4.3 and using the fact that we know $c_{m,n}$ for $m \leq 3$, $n \leq 8$. Similarly we can evaluate $c(45)$ by looking at the coefficient of r^5q^9. Notice that the term $(\sum_{m,n} c_{m,n} r^m q^n)^3/3$ has 3's only in the denominators of coefficients of $r^m q^n$ when $3|m$ and $3|n$, so we do not get problems from this for the coefficients of r^4q^9 and r^5q^9 (but we do get problems from this if we try to work out $c_g^-(27)$ using the same method). This proves lemma 4.4.

If $p = 13$ we run into trouble when we try to determine $c_g^-(26)$; this is why the argument in lemma 4.4 does not work in this case. Also the argument breaks down if we try to work out $c(27)$, because when we look at the coefficient of r^3q^9 we get an extra term $c_g^-(3)/3$ which does not have coefficients in $\mathbf{Z}_3$. This is why we use the coefficients $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ rather than $c_g^-(1)$, $c_g^-(2)$, $c_g^-(3)$, and $c_g^-(5)$.

Lemma 4.4 is the reason why our arguments do not work for $p < 13$ (and do not work so smoothly for $p = 13$). As p gets smaller we have fewer relations to work with and it gets more difficult to determine all coefficients in terms of the $c_g^-(m)$'s for small m. It is probably possible to extend lemma 4.4 to cover some smaller primes with a lot of effort. Fortunately it is not necessary to do the cases $p < 13$ because these cases have already been done by explicit calculations in [B-R].

Lemma 4.5. *If $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are known mod 3^n for some $n \geq 1$, and $p > 13$, and $(c_g^-(1), 3) = 1$ if $p = 13$, then these values of $c_g^-(m)$ are determined mod 3^{n+1}.*

Proof. By lemma 4.4 we see that we can determine the values mod 3^n of $c_g^-(n)$ for $n \leq 16$, $c_g^-(2n)$ for $2n \leq 32$, and $c_g^-(36)$ and $c_g^-(45)$.

But now if we look at the coefficient of r^3q^{3m} of 3.4 with $m = 1, 2, 4$, or 5, we see that $c_g^-(9m) + c_g^+(2m)/2 + c_g^-(m)/3 = c_g^-(m)^3/3+$ (some polynomial in known $c_g^-(i)$'s with coefficients that are 3-adic integers). But $c_g^-(m)^3/3$ mod 3^n depends only on $c_g^-(m)$ mod 3^n if $n \geq 1$, so we can determine $c_g^-(m)/3$ mod 3^n and hence $c_g^-(m)$ mod 3^{n+1}. This proves lemma 4.5.

This lemma is the reason that we use 3-adic rather than 2-adic approximation. If we try to prove the lemma above for 2^{n+1} instead of 3^{n+1}, all we find is that we can determine the $c_g^+(m)$'s mod 2^{n+1}, which is useless because we already know these numbers.

Proposition 4.6. *If the numbers $c_g^-(n)$ satisfy the conditions 1, 2, and 4 of lemma 4.2 then the numbers $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are congruent mod 3^{29} to the coefficients of q, q^2, q^4, and q^5 of one of the following power series. (The second column is a genus zero group whose Hauptmodul appears to be the function with coefficients $c_g^-(n)$. This has not been checked rigorously as it is not necessary for the proof of theorem 4.1.)*

$p = 71$	$\Gamma_0(71)+$	q^{-1}	$+q$	$+q^2$	$+q^3$	$+q^4$	$+2q^5$	$+\cdots$
$p = 59$	$\Gamma_0(59)+$	q^{-1}	$+q$	$+q^2$	$+2q^3$	$+2q^4$	$+3q^5$	$+\cdots$
$p = 47$	$\Gamma_0(47)+$	q^{-1}	$+q$	$+2q^2$	$+3q^3$	$+3q^4$	$+5q^5$	$+\cdots$
$p = 47$	$\Gamma_0(94)+$	q^{-1}	$+q$		$+q^3$	$+q^4$	$+q^5$	$+\cdots$
$p = 41$	$\Gamma_0(41)+$	q^{-1}	$+2q$	$+2q^2$	$+3q^3$	$+4q^4$	$+7q^5$	$+\cdots$
$p = 41$	$\Gamma_0(82\mid 2)+$	q^{-1}			$+q^3$		$+q^5$	$+\cdots$
$p = 31$	$\Gamma_0(31)+$	q^{-1}	$+3q$	$+3q^2$	$+6q^3$	$+9q^4$	$+13q^5$	$+\cdots$
$p = 31$	$\Gamma_0(62)+$	q^{-1}	$+q$	$+q^2$	$+2q^3$	$+q^4$	$+3q^5$	$+\cdots$
$p = 29$	$\Gamma_0(29)+$	q^{-1}	$+3q$	$+4q^2$	$+7q^3$	$+10q^4$	$+17q^5$	$+\cdots$
$p = 29$	$\Gamma_0(58\mid 2)+$	q^{-1}	$+q$		$+q^3$		$+q^5$	$+\cdots$
$p = 23$	$\Gamma_0(23)+$	q^{-1}	$+4q$	$+7q^2$	$+13q^3$	$+19q^4$	$+33q^5$	$+\cdots$
$p = 23$	$\Gamma_0(46) + 23$	q^{-1}		$-q^2$	$+q^3$	$-q^4$	$+q^5$	$+\cdots$
$p = 23$	$\Gamma_0(46)+$	q^{-1}	$+2q$	$+q^2$	$+3q^3$	$+3q^4$	$+5q^5$	$+\cdots$
$p = 19$	$\Gamma_0(19)+$	q^{-1}	$+6q$	$+10q^2$	$+21q^3$	$+36q^4$	$+61q^5$	$+\cdots$
$p = 19$	$\Gamma_0(38)+$	q^{-1}	$+2q$	$+2q^2$	$+5q^3$	$+4q^4$	$+9q^5$	$+\cdots$
$p = 19$	$\Gamma_0(38\mid 2)+$	q^{-1}	$+2q$		$+q^3$		$+3q^5$	$+\cdots$
$p = 17$	$\Gamma_0(17)+$	q^{-1}	$+7q$	$+14q^2$	$+29q^3$	$+50q^4$	$+92q^5$	$+\cdots$
$p = 17$	$\Gamma_0(34)+$	q^{-1}	$+3q$	$+2q^2$	$+5q^3$	$+6q^4$	$+12q^5$	$+\cdots$
$p = 17$	$\Gamma_0(34\mid 2)+$	q^{-1}	$+q$		$+3q^3$		$+4q^5$	$+\cdots$
$p = 13$	$\Gamma_0(13)+$	q^{-1}	$+12q$	$+28q^2$	$+66q^3$	$+132q^4$	$+258q^5$	$+\cdots$
$p = 13$	$\Gamma_0(26)+$	q^{-1}	$+4q$	$+4q^2$	$+10q^3$	$+12q^4$	$+26q^5$	$+\cdots$
$p = 13$	$\Gamma_0(26\mid 2)+$	q^{-1}	$+2q$		$+4q^3$		$+6q^5$	$+\cdots$

It is also possible (but unlikely) that there are other solutions for $p = 13$ for which $c_g^-(1)$ does not satisfy the inequalities $0 \leq c_g^-(1) < 3^{10}$.

Proof. For each prime p we use a computer to test all 81 possibilities for $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ mod 3. If $p > 13$ then we can calculate the p-adic expansion of all the coefficients $c_g^-(n)$ recursively using lemma 4.5 above, and we reach a contradiction by looking at coefficients of proposition 4.3 except in the cases above. For $p = 13$ this does not quite work as we have not shown the values mod 3^n determine those mod 3^{n+1}, so we can adopt the crude procedure of just testing all 3^4 possibilities for the $c_g^-(i)$'s mod 3^{n+1} for each solution mod 3^n we have found, and checking to see which of them leads to contradictions. There were some cases for $p = 13$ where this did not lead to a contradiction but did at least lead to the conclusion that $c_g^-(1) < 0$ or $c_g^-(1) \geq 3^{10}$. (If $c_g^-(1)$ is not divisible by 3 then even when $p = 13$ the numbers $c_g^-(i)$ mod 3^n determine the numbers $c_g^-(n)$ mod 3^{n+1}. When $c_g^-(1)$ was divisible by 3, there were several cases where the coefficient $c_g^-(5)$ was not determined mod 3^{n+1} by the identities used by the computer program. However in all the cases looked at by the computer with $3|c_g^-(1)$, the values of $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $3c_g^-(5)$ mod 3^n uniquely determined their values mod 3^{n+1}.) This proves proposition 4.6, at least if one believes the computer calculations. (Anyone who does not like computer calculations in a proof is welcome to redo the calculations by hand.)

Lemma 4.7. *If the coefficients $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are equal to the coefficients $c_g^+(1)$, $c_g^+(2)$, $c_g^+(4)$, and $c_g^+(5)$ for g an element of prime order $p > 13$ or an element of type $13A$ then all the coefficients are determined by the relations of proposition 3.4.*

Proof. This proof is just a long case by case check using induction. We will repeatedly use the fact that the coefficients of $r^m q^n$ of both sides of proposition 4.3 are equal.

Case 1. $c_g^-(2n)$ for $2n \not\equiv 0 \bmod p$. By looking at the coefficient of $r^2 q^n$ and using the fact that $c_{2,n} = c_{1,n+1}$ (as $(p, n) = 1$) we see that $c_g^-(2n)$ can be written as a polynomial in $c(n + 1)$ and $c(i)$ for $1 \le i < n$.

Case 2. $c_g^-(2n+1)$ for $2n+1 \not\equiv \pm 1 \bmod p$. By looking at the coefficient of $r^2 q^{2n}$ we can express $c_g^-(2n+1)$ in terms of $c_g^-(4n)$ and known quantities. But $c_g^-(4n)$ can be evaluated by looking at the coefficient of $r^4 q^n$ and using the fact that $c_{4,n} = c_{3,n+1} = c_{2,n+2}$ (as $2n + 1 \not\equiv \pm 1 \bmod p$).

The recursion formulas given by cases 1 or 2 can be written out explicitly and are given in [B, (9.1)]. In that paper they are sufficient to determine all coefficients because the analogues of the relations in proposition 3.4 hold whenever $mn \neq 0$.

The remaining cases of the proof are unusually tedious so we only sketch them briefly.

Case 3. $c_g^-(2n)$ for $2n \equiv 0 \bmod p$, n odd. We can express this in terms of $c_{2,n} = c_{3,n-1}$, which we can evaluate by calculating $c_g^-(3(n-1)) = c_g^-(2(3(n-1)/2))$ using case 1.

For the rest of the proof it is convenient to evaluate the cluster of elements $c_g^-(2pn - 1)$, $c_g^-(2pn + 1)$, and $c_g^-(4pn)$, simultaneously. (These are the only cases left to do.)

Case 4. $2np \equiv 2 \bmod 3$. We evaluate $c_g^-(2np + 1)$ in terms of $c_{3,(2np+1)/3}$. We evaluate $c_g^-(2np - 1)$ by expression $c_g^-(2np - 1)c(2)$ in terms of $c_{2,2np+1}$ and evaluating the latter by comparing it with $c_{3,2(2np+1)/3}$. We evaluate $c_g^-(4np)$ by expressing $c_{3,2np+2}$ in terms of $c_g^-(2)c_{2,2np}$ and evaluating the latter by comparing it with $c_{2,3np+3}$.

Case 5. $2np \equiv 1 \bmod 3$. We evaluate $c_g^-(2np-1)$ in terms of $c_{3,(2np-1)/3}$. We evaluate $c_{1,2np+5}$ by using $c_{3,(2np+5)/3}$ and then evaluate in turn $c_{2,2np+4} = c_{1,2np+5}$, $c_g^-(1)c_{1,2np+3}$ (using the coefficient of $r^2 q^{2np+4}$), $c_{2,2np+3} = c_{1,2np+3}$, $c_g^-(1)c_{1,2np+1}$ which gives us the value of $c_{1,2np+1}$. We evaluate $c_{2,2np}$ as in case 4.

Case 6. $2np \equiv 0 \bmod 3$. By looking at $c_{2,2np+2}$ and $c_{2,2np+3}$ we can find expressions for the linear combinations $c_g^-(1)c_g^-(2np + 1) + c_g^-(3)c_g^-(2np - 1)$ and $c_g^-(2)c_g^-(2np + 1) + c_g^-(4)c_g^-(2np - 1)$. If $p \neq 59$ or 71 then these two linear equations are independent so we can solve for $c_g^-(2np + 1)$ and $c_g^-(2np - 1)$, and then find $c_{2,2np}$ as in case 4. If $p = 59$ or 71 then an even more tedious argument using $c_{2,2np+2}$, $c_{3,2np+2}$, and $c_{3,2np+3}$ produces 3 independent linear relations between $c_g^-(2pn - 1)$, $c_g^-(2pn + 1)$, and $c_g^-(4pn)$, so we can solve for them.

This proves lemma 4.7.

We can now complete the proof of theorem 4.1. If we look at the possible solutions for the coefficients $c_g^-(n)$ given by proposition 4.6, we see that all the solutions except those congruent mod 3^{29} to the Hauptmoduls of the elements $g \in M$ are ruled out by the inequalities of lemma 4.2. As $3^{29} > c_1^+(i) > c_g^-(i) \ge 0$ for $i \le 5$ we see that $c_g^+(i) = c_g^-(i)$ for $i \le 5$. The coefficients $c_g^+(n)$ also satisfy the identities of proposition 4.3 satisfied by the $c_g^-(n)$'s; for example, this follows easily from the identity

$$\sum_{n\in\mathbf{Z}} c_g^+(n)r^n - \sum_{n\in\mathbf{Z}} c_g^+(n)q^n = r^{-1} \prod_{m>0,n\in\mathbf{Z}} (1 - r^m q^n)^{c_g^+(mn)}(1 - r^{pm}q^{pn})^{c_g^+(pmn)}$$

[B, page 434]. Therefore $c_g^-(n) = c_g^+(n)$ for all n by lemma 4.7 because they both have the same values for $n \le 5$ and both satisfy the same recursion relations given in proposition 3.4. This proves theorem 4.1.

5. Calculation of some cohomology groups.

In this section we calculate the Tate cohomology groups of some spaces that appear in the construction of the monster vertex algebra. In the next section we will use these calculations to find the Tate cohomology $\hat{H}^*(g, V[1/2])$ for elements g of types $3B$, $5B$, $7B$ and $13B$.

We will write $g_0, g_1, \ldots, g_{p-1}$ for the elements $1, g, \ldots, g^{p-1}$ of the group ring $\mathbf{Z}_p[G]$ to avoid confusing the group ring multiplication with the multiplication in other rings containing it.

Suppose that A is a free $\mathbf{Z}_p$-module. We write $h(A)$ for the $\mathbf{Z}$-graded $\mathbf{Z}_p$-algebra generated by elements $h_n(a)$ of degree n for $a \in A$, $n \geq 0$ with the relations

1. $h_n(\lambda a) = \lambda^n h_n(a)$ if $\lambda \in \mathbf{Z}_p$.
2. $h_0(a) = 1$, $h_n(a) = 0$ if $n < 0$.
3. $h_n(a+b) = \sum_{i \in \mathbf{Z}} h_i(a) h_{n-i}(b)$.

If A has a basis $a_1, \ldots, a_m$ then $h(A)$ is the polynomial algebra generated by elements $h_n(a_i)$ for $n > 0$, $1 \leq i \leq m$. An action of G on A induces an action of G on $h(A)$ in the obvious way.

Lemma 5.1. *The cohomology ring $\hat{H}^*(g, h(\mathbf{Z}_p[G]))$ is the polynomial algebra over $\mathbf{F}_p$ generated by the ordinary elements $h_n(g_0) + h_n(g_1) + \cdots + h_n(g_{p-1})$ for $n > 0$.*

Proof. This follows from the fact that $h(\mathbf{Z}_p[G])$ is the tensor product of the rings A_n, where A_n is the polynomial algebra generated by the elements $h_n(g_i)$ for $0 \leq i < p$. We see that A_n is isomorphic to the symmetric algebra $S^*(\mathbf{Z}_p[G])$ whose cohomology is given by lemma 2.8, and we then apply corollary 2.2 to finish the proof. (Notice that the space spanned by the elements $h_n(g_i)$ for $0 \leq i < p$ is not $h_n(\mathbf{Z}_p[G])$; in fact, the latter is not even a $\mathbf{Z}_p$-module.) This proves lemma 5.1.

Proposition 5.2. *The cohomology ring $\hat{H}^*(g, h(I))$ is isomorphic to the free exterior algebra with one super generator in each positive degree not divisible by p. The automorphism -1 of I multiplies each generator by -1 and hence acts as -1 on $\hat{H}^1(g, h(I))$ and as 1 on $\hat{H}^0(g, h(I))$.*

Proof. The algebra $h(I)$ is the quotient of $h(\mathbf{Z}_p[G])$ by the ideal generated by the elements $h_n(N_G)$. We let R_m be the quotient of $h(\mathbf{Z}_p[G])$ by the elements $h_n(N_G)$ for $n \leq m$, and we will prove by induction on m that $\hat{H}^*(R_m)$ is the tensor product of the free polynomial algebra generated by the ordinary elements $h_n(g_0) + h_n(g_1) + \cdots + h_n(g_{p-1})$ (of degree n) for $n > m/p$ and the free exterior algebra generated by super elements of each degree n with $0 < n \leq m$, $(p, n) = 1$ (of degree n).

To do this we look at the exact sequence

$$0 \to R_{m-1} \to R_{m-1} \to R_m \to 0$$

where the map from R_{m-1} to R_{m-1} is given by multiplication by $h_m(N_G)$, and work out the cohomology of R_m using the cohomology exact sequence of this short exact sequence. We have to consider the cases $p|m$ and $(p, m) = 1$ separately.

If $(p, m) = 1$ then multiplication by $h_m(N_G)$ induces the zero map on cohomology, so that $\hat{H}^*(g, R_m)$ is the sum of $\hat{H}^*(g, R_{m-1})$ and $\hat{H}^*(g, R_{m-1})$ with the degree shifted by m and with the ordinary and super parts exchanged. In other words, $\hat{H}^*(g, R_m)$ is the tensor product of $\hat{H}^*(g, R_{m-1})$ with an exterior algebra generated by one super element of degree m. This proves the induction step when $(p, m) = 1$.

If $p|m$ then multiplication by $h_m(N_G)$ induces an injective map on cohomology, so that $\hat{H}^*(g, R_m)$ is the quotient of $\hat{H}^*(g, R_{m-1})$ by the image of $h_m(N_G)$, in other words it is just $\hat{H}^*(g, R_{m-1})$ with a polynomial generator of degree m killed off.

Hence we see that to go from $\hat{H}^*(h(R_{m-1}))$ to $\hat{H}^*(h(R_m))$ we either kill off a polynomial generator or add an exterior algebra generator. This proves the statement about $\hat{H}^*(h(R_m))$ for all m by induction, and hence proves proposition 5.2.

For completeness we remark that the cohomology ring $\hat{H}^*(g, h(\mathbf{Z}_p))$ is a polynomial algebra over $\mathbf{F}_p$ with one ordinary generator of each positive degree.

For any $\mathbf{Z}_p$-module A we define ω to be the automorphism of $h(A)$ taking $h_n(a)$ to $(-1)^n h_n(-a)$. We define $h_\omega(A)$ to be the largest quotient ring of $h(A)$ on which the automorphism ω acts trivially. (Notice that this is smaller than the largest quotient module on which ω acts trivially.)

Lemma 5.3. *The cohomology ring $\hat{H}^*(g, h_\omega(\mathbf{Z}_p[G]))$ is the polynomial ring on ordinary generators of degrees $(2n+1)p$ for $n \geq 0$.*

Proof. The algebra $h_\omega(\mathbf{Z}_p[G])$ is the polynomial algebra generated by the elements $h_{2n+1}(g_i)$ for $0 \leq i < p$, $n \geq 0$, and lemma 5.3 follows from this.

Lemma 5.4. *The cohomology ring $\hat{H}^*(g, h_\omega(I))$ is the exterior algebra on generators of degrees $(2n+1)$ for $n \geq 0$, $(n,p) = 1$. In particular it is generated (as a ring) by super elements of odd degrees, so that $\hat{H}^0(g, h_\omega(I))$ vanishes in odd degrees and $\hat{H}^1(g, h_\omega(I))$ vanishes in even degrees.*

This follows from lemma 5.3 in the same way that lemma 5.2 follows from 5.1.

Remark. The ring $h(\mathbf{Z}_p)$ can be identified with the ring of symmetric functions over $\mathbf{Z}_p$ if we identify $h_n(1)$ with the n'th complete symmetric function. In this case the automorphism ω taking $h_n(1)$ to $(-1)^n h_n(-1)$ is just the automorphism ω of [Mac, p. 14] taking the complete symmetric functions to the elementary symmetric functions.

Remark. More generally we can form a quotient ring associated to any automorphism ω of A of order n prime to p such that $h_\omega(A)$ is the quotient ring associated to the automorphism $\omega = -1$ of A. To do this we let R_p be the unramified extension of $\mathbf{Z}_p$ generated by a primitive n'th root of unity ζ. We then form the largest quotient ring of $h(A) \otimes R_p$ whose elements are fixed by the automorphism taking $h_n(a)$ to $\zeta^n h_n(\omega(a))$. For example this space appears when we try to construct the monster vertex algebra from a fixed point free automorphism ω of the Leech lattice Λ.

The extraspecial group 2^{1+24} occurring in the centralizer of an element of type $2B$ in M has a unique 2^{12}-dimensional irreducible module A over $\mathbf{Z}[1/2]$. Any element of odd order in $\mathrm{Aut}(\Lambda)$ has an induced action on 2^{1+24} and on A.

Lemma 5.5. *If g is an element of order 3, 5, 7, or 13 acting on the Leech lattice Λ with no nonzero fixed vectors then as a $\mathbf{Z}_p[G]$-module, $A \otimes \mathbf{Z}_p$ is the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$.*

Proof. We let B be the 2^{24}-dimensional 2^{1+24}-module $A \otimes A$. This is isomorphic to the group ring of 2^{24}. The element g acts fixed point freely on the group $2^{24} = \Lambda/2\Lambda$ because it acts fixed point freely on Λ and has odd order, so as a module over G, $B \otimes \mathbf{Z}_p$ is the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$, and the Tate cohomology groups are therefore $\hat{H}^0(g, B \otimes \mathbf{Z}_p) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g, B \otimes \mathbf{Z}_p) = 0$. As $\hat{H}^*(g, B \otimes \mathbf{Z}_p) = \hat{H}^*(g, A \otimes \mathbf{Z}_p) \otimes \hat{H}^*(g, A \otimes \mathbf{Z}_p)$ we see that $\hat{H}^*(g, A \otimes \mathbf{Z}_p)$ is either an ordinary or a super vector space over $\mathbf{F}_p$ of dimension 1.

Therefore as a G module, $A \otimes \mathbf{Z}_p$ is the sum of some copies of $\mathbf{Z}_p[G]$ and either $\mathbf{Z}_p$ or I. But $\dim(A) = 2^{12} \equiv 1 \bmod p$ as $(p-1)|12$ which rules out the second possibility. Therefore $\hat{H}^0(g, A \otimes \mathbf{Z}_p)) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g, A \otimes \mathbf{Z}_p) = 0$. This proves lemma 5.5.

A similar argument can be used to calculate $\hat{H}^*(g, A)$ for any element g of odd prime order p in $\mathrm{Aut}(\Lambda)$. If $2n$ is the dimension of the subspace of Λ fixed by g then n is an integer with $2^n \equiv \pm 1 \bmod p$. Exactly one of the two cohomology groups $\hat{H}^*(g, A)$ vanishes and the other has dimension 2^n over $\mathbf{F}_p$. The group $\hat{H}^1(g, A)$ vanishes if $2^n \equiv 1 \bmod p$, and the group $\hat{H}^0(g, A)$ vanishes if $2^n \equiv -1 \bmod p$. There are some cases where it is $\hat{H}^0(g, A)$ and not $\hat{H}^1(g, A)$ that vanishes; for example, if g is the element of order 5 with trace -1 on Λ then $\hat{H}^0(g, A) = 0$ and $\hat{H}^1(g, A) = (\mathbf{Z}/5\mathbf{Z})^4$.

6. Elements of type $3B$, $5B$, $7B$, and $13B$.

We use the result of the previous section to calculate the cohomology groups $\hat{H}^i(g, V[1/2])$ when $g \in M$ has type $3B$, $5B$, $7B$, or $13B$.

Theorem 6.1. *If $g \in M$ has type $3B$, $5B$, $7B$, or $13B$ and σ is the element of order 2 generating the center of $C_M(g)/O_p(C_M(g))$ then $O_p(C_M(g))$ acts trivially on ${}^gV$ and σ fixes the ordinary part of ${}^gV$ and acts as -1 on the super part.*

Proof. We prove this by calculating the action of σ on $\hat{H}^*(g, V[1/2])$ explicitly. We recall from [B-R] that the $\mathbf{Z}[1/2]$-form $V[1/2]$ of Frenkel, Lepowsky and Meurman's monster vertex algebra [FLM] is the sum of two pieces V^0 and V^1 which are the $+1$ and -1 eigenspaces of an element σ of type $2B$. (Unfortunately the details for the construction of this $\mathbf{Z}[1/2]$-form have not been published anywhere.) We now show that $\hat{H}^0(g, V^1) = \hat{H}^1(g, V^0) = 0$.

The space V^0 is the subspace of $V_\Lambda[1/2]$ fixed by the automorphism induced by -1 acting on the Leech lattice Λ, where $V_\Lambda[1/2] = V_\Lambda \otimes \mathbf{Z}[1/2]$, and V_Λ is the integral form of the vertex algebra of the Leech lattice, described in [B-R]. As in [B-R], the space V_Λ is the tensor product of a twisted group ring $\mathbf{Z}[\hat{\Lambda}]$ and a space isomorphic to $h(\Lambda)$. The cohomology $\hat{H}^*(g, \mathbf{Z}[\hat{\Lambda}])$ is just $\mathbf{Z}/p\mathbf{Z}$ as in [B-R] because g acts fixed point freely on Λ. By proposition 5.2, the element -1 of $\mathrm{Aut}(\hat{\Lambda})$ acts as $+1$ on $\hat{H}^0(g,h(\Lambda))$ and as -1 on $\hat{H}^1(g,h(\Lambda))$. Therefore $\hat{H}^1(g,V^0)$ vanishes because it is isomorphic to the subspace of $\hat{H}^1(g,h(\Lambda))$ fixed by $-1\in\mathrm{Aut}(\Lambda)$.

The space V^1 is the subspace of elements of integral weight of the tensor product $A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p)$, where A is the 2^{12} dimensional module for 2^{1+24} and $h_\omega(\Lambda\otimes\mathbf{Z}_p)$ has its grading halved and then shifted by 3/2. By lemma 5.5, $\hat{H}^*(g,A)=\mathbf{F}_p$, so $\hat{H}^*(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))=\hat{H}^*(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$. By lemma 5.4 the cohomology ring $\hat{H}^*(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ is generated by super elements of odd degree, so $\hat{H}^1(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in even degrees and $\hat{H}^1(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in odd degrees. Therefore $\hat{H}^0(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in integral degrees, so $\hat{H}^0(g,V^1)=0$.

This shows that the element σ acts as 1 on $\hat{H}^0(g,V[1/2])=\hat{H}^0(g,V_\Lambda)=\hat{H}^0(g,h(\Lambda\otimes\mathbf{Z}_p))$ and as -1 on $\hat{H}^1(g,V[1/2])=\hat{H}^1(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))$.

Finally we show that $O_p(C_M(g))$ acts trivially on ${}^gV$. The element σ acts trivially on the center $\langle g\rangle$ of the extraspecial group $p^{1+24/(p-1)}$ and acts as -1 on the quotient $p^{24/(p-1)}$. The element g acts trivially on ${}^gV$ so we get an action of $p^{24/(p-1)}$ on ${}^gV$. As σ acts as 1 on the even part of ${}^gV$ and as -1 on the odd part, it commutes with all automorphisms of ${}^gV$, so the group $p^{24/(p-1)}$ must act trivially on ${}^gV$. This proves theorem 6.1.

The group $C_M(g)$ has structure $3^{1+12}.2.Suz$, $5^{1+6}.2.HJ$, $7^{1+4}.2.A_7$, or $13^{1+2}.2.A_4$, so we get modular vertex superalgebras acted on faithfully by the groups $2.Suz$, $2.HJ$, $2.A_7$, and $2.A_4$ whose modular characters are given by Hauptmoduls. We can calculate the characters of both the ordinary and super parts of these superalgebras by using the fact that the characters of their sum and difference are both given in terms of known Hauptmoduls.

It seems likely that if g has type $2B$ then the vertex superalgebra acted on by $2^{24}.Co_1$ constructed in [B-R] is acted on trivially by the 2^{24} subgroup, but this has not yet been proved. The proof given above for the corresponding fact for elements of odd prime order fails for order 2 because an automorphism mapping each element of 2^{24} to its inverse is still the identity automorphism.

As an example, we work out the dimensions of the ordinary and super parts of ${}^gV$ in the case when g has type $3B$. The function $\sum_n\mathrm{Tr}(1|{}^gV)$ is equal to the Hauptmodul of the element g of type $3B$ in the monster, which is

$$q^{-1}+54q-76q^2-243q^3+1188q^4-1384q^5-\cdots,$$

and the function $\sum_n\mathrm{Tr}(\sigma|{}^gV)$ is equal to the Hauptmodul of the element σg of type $6B$ in the monster, which is

$$q^{-1}+78q+364q^2+1365q^3+4380q^4+12520q^5+\cdots$$

so we find that the dimension of the ordinary part of ${}^gV_n$ is the coefficient of q^n of

$$q^{-1}+66q+144q^2+561q^3+2784q^4+5568q^5+\cdots$$

and the dimension of the super part of ${}^gV_n$ is the coefficient of q^n of

$$12q+220q^2+804q^3+1596q^4+6952q^5\cdots.$$

These series can be written as sums of dimensions of ordinary representations of the groups $3.Suz$ and $6.Suz$ as

$$q^{-1}+66q+(1+143)q^2+(429+66+66)q^3+(66+66+78+429+2145)q^4+(4\times1+4\times143+2\times780+3432)q^5+\cdots$$

and

$$12q+220q^2+(780+12+12)q^3+(780+780+12+12+12)q^4+(4\times220+2\times572+4928)q^5+\cdots.$$

(Some of the decompositions are ambiguous; for example $1+143$ could be $66+78$. We have chosen $1+143$ instead of $66+78$ for a reason that will appear in a moment.) At first sight this seems to rule out a vertex algebra structure on any lift to characteristic 0, as the homogeneous components V_n, $V_{n+1},\ldots$ have to increase. However we can get around this by rescaling the grading by a factor of 1/3. The data above seem consistent with the existence of a $\frac{1}{3}\mathbf{Z}$-graded vertex superalgebra $W=\sum_{3n\in\mathbf{Z}}W_n$ such that the mod

3 reduction of W_n is V_{3n}. The vertex superalgebra W is probably acted on by $6.Suz$, with the element of order 3 in the center acting as ω^{3n} on W_n.

There is a second modular vertex super algebra acted on by each of the groups $2.Suz$, $2.HJ$, $2.A_7$, and $2.A_4$, which can be constructed as $\hat{H}^*(g, V_\Lambda)$ where g is a fixed point free automorphism of order 3, 5, 7, or 13 of the Leech lattice Λ. The ordinary part of this superalgebra is the same as the ordinary part of the superalgebra constructed from the monster vertex algebra, but its super part is different; for example, its degree 0 piece does not vanish.

7. Open problems and conjectures.

For more open problems in this area see section 7 of [B-R].

At the end of section 6 we saw that the vertex superalgebra of an element of type $3B$ looks as if it might be the reduction mod 3 of a superalgebra in characteristic 0. More generally something similar seems to happen for some other elements of the monster, as described in the following conjecture.

Conjecture. *For each element $g \in M$ of order n there is a $\frac{1}{n}\mathbf{Z}$-graded super module ${}^g\hat{V} = \sum_{i \in \frac{1}{n}\mathbf{Z}} V_i$ over the ring $\mathbf{Z}[e^{2\pi i/n}]$ with the following properties.*

1. *${}^g\hat{V}$ is often a vertex superalgebra with a conformal vector and a self dual invariant symmetric bilinear form. (For example, it seems to be a vertex algebra when g is of type $1A$, $2B$, $3B$, $5B$, $7B$, $13B$, or $3C$, but possibly not if g has type $2A$ or $3A$.)*
2. *${}^1\hat{V}$ is a self dual integral form of the monster vertex algebra (probably the one conjecturally described in lemma 7.1 of [B-R]).*
3. *If g has prime order then the reduction mod p of ${}^g\hat{V}$ is the modular vertex superalgebra ${}^gV$ associated to g (at least when its grading is multiplied by n.)*
4. *If all coefficients of the Hauptmodul of g are nonnegative then the super part of ${}^g\hat{V}$ vanishes. In this case ${}^g\hat{V}$ should be the same as the "twisted sector" of g that appears in Norton's generalized moonshine conjectures [N]. (But notice that if the Hauptmodul of g has negative coefficients then ${}^g\hat{V}$ is certainly not the twisted sector of g; for example, ${}^g\hat{V}$ is a super vector space and has a nontrivial piece of degree 0.)*
5. *${}^g\hat{V}$ is acted on by a central extension $(\mathbf{Z}/n\mathbf{Z}).C_M(g)$ of $C_M(g)$ by $\mathbf{Z}/n\mathbf{Z}$. This central extension contains an element of order 1 or 2 acting as -1 on the super part of ${}^g\hat{V}$ and as 1 on the ordinary part.*
6. *If g and h are commuting elements of M of coprime orders and $\hat{g}$ is the lift of g to the central extension of $C_M(h)$ with the same order as g, then*
$$\mathrm{Tr}(\hat{g}|{}^h\hat{V}) = \mathrm{Tr}(gh|{}^1\hat{V}).$$
(This identity is often false if the orders of g and h are not coprime.)
7. *If g and h commute then $\hat{H}^*(\hat{g}, {}^h\hat{V})$ is (essentially) the reduction mod $|g|$ of ${}^{gh}V$. (Actually this is not quite correct as stated, because the reduction mod $|g|$ will be defined over some nilpotent extension of $\mathbf{Z}/|g|\mathbf{Z}$, and we should then take the tensor product over this ring with $\mathbf{Z}/|g|\mathbf{Z}$.)*

These conjectures are similar to Norton's generalized moonshine conjectures [N] because both associate some graded space with each element g of the monster. However they treat the problem of negative coefficients in the Hauptmodul $T_g(\tau)$ in a different way: Norton's conjectures use the fact that $T_g(-1/\tau)$ has nonnegative coefficients, which are supposed to be the dimensions of some vector spaces, while the conjectures above allow super vector spaces which may have negative super dimension. Perhaps both Norton's conjectures and the conjectures above are special cases of some conjecture which associates a $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}$-graded super space acted on by a central extension $\mathbf{Z}/n\mathbf{Z}.C_M(g)$ to each element $g \in M$ of order n.

We finish by listing a few other questions.

1. Is the group $C_M(g)/O_p(C_M(g))$ the full automorphism group of the modular vertex algebra ${}^gV$? If g has type $3C$ it seems possible that it is not, and the full automorphism group may be $E_8(\mathbf{F}_3)$. Notice that if g has type $2B$ then it is not yet known that $O_2(C_M(g))$ acts trivially on ${}^gV$.
2. Replace the revolting argument of section 4 that involves computer calculations with a more conceptual proof. More generally, suppose we have coefficients $c(n)$ of some function such that the coefficient of r^mq^n in
$$r^{-1}\prod_{m>0, n\in\mathbf{Z}}(1-r^mq^n)^{c(mn)}$$

vanishes whenever $(mn, N) = 1$ for some fixed integer N. What does this imply about the coefficients $c(n)$? (There are many solutions to this given by taking the $c(n)$'s to be the coefficients of some Hauptmodul of level N.)

3. What happens when we apply the argument of section 3 to some automorphism of prime order of the fake monster Lie algebra? Do the infinite products we get represent interesting functions?
4. In [B section 10] it was suggested that there should be Lie superalgebras associated with elements g of the monster, but it was not clear what the dimensions of homogeneous components of these superalgebras should be in the cases when the coefficients of the Hauptmodul of g were neither all positive nor alternating in sign. The results of section 6 suggest what to do in this case: the sum of the dimensions of the ordinary and super parts should also be given by the coefficients of another Hauptmodul.

References.

[BBCO] C. Batut, D. Bernardi, H. Cohen, M. Olivier, "User's guide to PARI-GP". This guide and the PARI programs can be obtained by anonymous ftp from hensel.mathp7.jussieu.fr in the directory /dist/pari.

[B] R. E. Borcherds, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math. 109, 405-444 (1992).

[B-R] R. E. Borcherds, A. J. E. Ryba, Modular Moonshine II, Duke Math Journal Vol. 83 No. 2, 435-459, 1996.

[CG] C. J. Cummins, T. Gannon, Modular equations and the genus zero property of moonshine functions, Invent. Math. 129, 413-443 1997.

[DM] C. Dong and G. Mason, On the construction of the moonshine module as a $\mathbf{Z}_p$-orbifold, Mathematical aspects of conformal and topological field theories and quantum groups (South Hadley, MA, 1992), 37–52, Contemp. Math., 175, Amer. Math. Soc., Providence, RI, 1994.

[FLM] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[Mac] I. G. Macdonald, "Symmetric functions and Hall polynomials", Oxford University press, 1979.

[Ma] Y. Martin, On modular invariance of completely replicable functions. Moonshine, the Monster, and related topics (South Hadley, MA, 1994), 263–286, Contemp. Math., 193, Amer. Math. Soc., Providence, RI, 1996.

[M] P. Montague, Third and Higher Order NFPA Twisted Constructions of Conformal Field Theories from Lattices, Nuclear Phys. B 441 (1995), no. 1-2, 337–382.

[N] S. P. Norton, Generalized Moonshine, Proc. Symp. Pure Math. 47 (1987) p. 208-209.

[R] A. J. E. Ryba, Modular Moonshine?, In "Moonshine, the Monster, and related topics", edited by Chongying Dong and Geoffrey Mason. Contemporary Mathematics, 193. American Mathematical Society, Providence, RI, 1996. 307-336.

arXiv:hep-th/9607026v1 3 Jul 1996

July 1996
CERN-TH/96-170

Counting Dyons in $N = 4$ String Theory

Robbert Dijkgraaf

Department of Mathematics
University of Amsterdam, 1018 TV Amsterdam

Erik Verlinde

TH-Division, CERN, CH-1211 Geneva 23
and
Institute for Theoretical Physics
Universtity of Utrecht, 3508 TA Utrecht

and

Herman Verlinde

Institute for Theoretical Physics
University of Amsterdam, 1018 XE Amsterdam
and
Joseph Henry Laboratories
Princeton University, Princeton, NJ 08544

Abstract

We present a microscopic index formula for the degeneracy of dyons in four-dimensional $N = 4$ string theory. This counting formula is manifestly symmetric under the duality group, and its asymptotic growth reproduces the macroscopic Bekenstein-Hawking entropy. We give a derivation of this result in terms of the type II five-brane compactified on $K3$, by assuming that its fluctuations are described by a closed string theory on its world-volume. We find that the degeneracies are given in terms of the denominator of a generalized super Kac-Moody algebra. We also discuss the correspondence of this result with the counting of D-brane states.

1. INTRODUCTION

In this paper we will study the dyonic spectrum of $N=4$ string theory in 4 dimensions. This theory has two perturbative formulations: one in terms of the toroidally compactified heterotic string, and a dual formulation in terms of type II strings compactified on $K3 \times T^2$. To describe the set of charged states we will take the point of view of the heterotic string, so that we have 28 electric charges q_e and 28 magnetic charges q_m which both lie on the $\Gamma^{22,6}$ lattice. This theory is conjectured [1, 2, 3] to have as its exact duality group

$$SL(2,\mathbf{Z}) \times SO(22,6,\mathbf{Z}) \tag{1.1}$$

with $SL(2,\mathbf{Z})$ being the electric-magnetic duality symmetry. The perturbative heterotic string states carry only electric charge, so that states with magnetic charge arise necessarily non-perturbatively as solitons.

Most accessible to computations is the spectrum of BPS states. These states are annihilated by some subset of the 16 supersymmetry charges and therefore many of their properties are protected in perturbation theory. The purely electric states are special in the sense that they preserve 1/2 of the supercharges and their degeneracies are easily determined, since they simply correspond to the heterotic string states that are in the right-moving ground state. Hence a BPS heterotic string state is described by specifying the 28 charges $q_e \in \Gamma^{22,6}$ together with the occupation numbers N_l^I ($I = 1, \ldots, 24$, $l > 0$), subject to the level-matching condition

$$\tfrac{1}{2} q_e^2 + \sum_{l,I} l N_l^I = 1. \tag{1.2}$$

Here the subscript l to N_l^I denotes the world-sheet oscillation number of the coordinate field X^I, and q_e^2 is defined using the $SO(22,6)$ invariant inner product on the $\Gamma^{22,6}$ lattice. The number of such states is given by[*]

$$d(q_e) = \oint d\sigma \, \frac{e^{i\pi\sigma q_e^2}}{\eta(\sigma)^{24}} \tag{1.3}$$

where the 'contour' integral over σ is from 0 to 1 and $\eta(\sigma)$ is the Dedekind η-function.

The conjectured electric-magnetic $SL(2,\mathbf{Z})$-duality predicts that there should also exist a solitonic version of the heterotic string that carries pure magnetic charge $q_m \in \Gamma^{22,6}$, and thus that a similar formula counts the pure magnetically charged 1/2 BPS states. The generic dyonic states, however, preserve only 1/4 of the supersymmetries and are more mysterious. In the following we will give a concrete proposal for the exact degeneracy of these dyonic BPS states. To be more precise, we will present a formula for the number of bosonic minus the number of fermionic BPS-multiplets for a given electric and magnetic charge. The basic idea is the following. We like to think of the dyonic states as some kind of bound state of an electric heterotic string with a dual magnetic heterotic string. The formula that counts these states, should therefore be a suitable generalization of (1.3), that will contain (1.3) as a special subcase. In addition it has to satisfy other non-trivial consistency checks, such as invariance under the full duality group, while it also has to reproduce the correct asymptotic growth as predicted from the macroscopic entropy formulas of extremal dyonic black holes [4, 5].

---

[*] Here and in the subsequent we omit a factor of 16, and therefore count the number of 1/2 BPS multiplets rather than individual states.

After presenting our formula and illustrating some of its features, we will show that the complete duality invariant dyon spectrum has a unified and natural interpretation in terms of strings on the five-brane. We will also discuss the correspondence with the counting of BPS-states that follows from a D-brane description [6, 7, 8, 9].

2. The degeneracy formula

To write the proposed dyonic index formula, it will be convenient to combine the electric and magnetic charge vectors q_e and q_m into a vector as follows

$$\boldsymbol{q} = \begin{pmatrix} q_m \\ q_e \end{pmatrix} \tag{2.1}$$

Accordingly we introduce the 2×2 matrix

$$\Omega = \begin{pmatrix} \rho & \upsilon \\ \upsilon & \sigma \end{pmatrix} \tag{2.2}$$

generalizing the single modulus σ in (1.3). The degeneracy formula will then take the following form

$$d(q_e, q_m) = \oint d\Omega \, \frac{e^{i\pi \boldsymbol{q}\cdot\Omega\cdot\boldsymbol{q}}}{\Phi(\Omega)}. \tag{2.3}$$

The integrals over the moduli parameters σ, ρ and υ again all run over the domain from 0 to 1, and impose level matching conditions analogous to (1.2).

In the above suggestive form, it is natural to identify the matrix Ω with the period matrix of a genus two Riemann surface. In the result that we will obtain, the denominator $\Phi(\Omega)$ indeed will turn out to be a genus two modular form. More precisely, it is the unique automorphic form of weight 10 of the modular group $Sp(2, \mathbf{Z})$ (represented by

4×4 matrices) and can be expressed as the squared product of all genus 2 theta functions with even spin structure

$$\Phi(\Omega) = 2^{-12} \prod_{\alpha=even} \theta[\alpha](\Omega)^2. \tag{2.4}$$

$\Phi(\Omega)$ also happens to be equal to the denominator in the Weyl-Kac-Borcherds character formula for a generalized super Kac-Moody algebra, and as such it is a special case of the automorphic forms constructed in [10]. The $SL(2,\mathbf{Z})$ duality transformations are identified with the subgroup of $Sp(2,\mathbf{Z})$ that leave the genus two modular form $\Phi(\Omega)$ invariant, and thus the presented degeneracies are manifestly duality symmetric. In addition, the formula satisfies a number of other non-trivial consistency checks, that we will now discuss. Subsequently, we will turn to the derivation and more detailed description of this result.

To make our formula somewhat more explicit, we can expand the integrand as a formal Fourier series expansion

$$\frac{1}{\Phi(\Omega)} = \sum_{k,l,m} D(k,l,m) e^{-2\pi i(k\rho + l\sigma + m\upsilon)} \tag{2.5}$$

with k, l, m integers. The coefficients $D(k,l,m)$ are all integers, that via (2.3) give us the degeneracy for a given electric and magnetic charge as

$$d(q_m, q_e) = D(\tfrac{1}{2}q_m^2, \tfrac{1}{2}q_e^2, q_e \cdot q_m) \tag{2.6}$$

Note that the normalization of (2.4) is chosen such that $d(0,0) = 1$.

Correspondence with the heterotic string

As we will explain in the following, the parameter υ couples to the helicity m of the dyonic states. The integral over υ thus projects on dyons with helicity equal to zero. However, instead of integrating over υ we can also put it to fixed value, like $\upsilon = 0$. The resulting formula will then correspond to a helicity trace $(-1)^m$, which will project out all 1/4 BPS representations, except for the purely electric or magnetic 1/2 BPS representations. Hence, for $\upsilon \to 0$ the integrand in formula (2.3) should match with the degeneracy formula (1.3) for the heterotic string BPS spectrum.

To make this correspondence explicit, we note that the function $1/\Phi(\Omega)$ is in fact identical to the chiral half of the bosonic string partition function (or rather, that of 24 massless scalars) on the corresponding genus two curve (see [11]). This means in particular that, if we factorize the genus two surface parametrized by Ω into two separate genus one surfaces with moduli ρ and σ, we indeed recover the separate 1/2 BPS partition sums. In

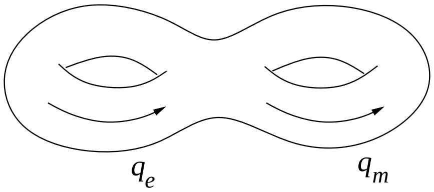


Fig. 1: *The dyon counting can be represented in terms of a genus two partition sum. The electric and magnetic charges of the dyon correspond to the loop momenta through the two handles of the Riemann surface.*

this analogy the electric and magnetic charges correspond to the loop momenta through the two handles, as in *fig.* 1.

Concretely, this factorization corresponds to taking the limit $\upsilon \to 0$, under which the function $1/\Phi(\Omega)$ reduces to

$$\frac{e^{i\pi \boldsymbol{q}\cdot\Omega\cdot\boldsymbol{q}}}{\Phi(\Omega)} \quad\longrightarrow\quad \frac{1}{\upsilon^2}\,\frac{e^{i\pi\rho q_m^2}}{\eta(\rho)^{24}}\,\frac{e^{i\pi\sigma q_e^2}}{\eta(\sigma)^{24}} \tag{2.7}$$

The diverging factor $1/\upsilon^2$ corresponds to the 'tachyon pole,' and reflects the emergence of two non-compact translational zero-modes for $\upsilon = 0$. In the other factors we recover the separate electric and magnetic heterotic string contributions. The above correspondence is a necessary boundary condition on any (candidate) dyonic degeneracy formula.

Asymptotic growth and black hole entropy

As a further non-trivial boundary condition, the formula (2.3) also matches the asymptotic behavior for large charges that one obtaines by comparing with the results for the macroscopic Bekenstein-Hawking entropy of extremal 4-dimensional black holes. This correspondence predicts that the statistical entropy of states with charge (q_e, q_m) is given by [5]

$$S = \pi\sqrt{q_e^2 q_m^2 - (q_e\cdot q_m)^2}. \tag{2.8}$$

The asymptotic growth of the degeneracies $d(q_e, q_m)$ indeed agrees with this predicted entropy, as it can be shown that in the large charge limit

$$d(q_e, q_m) \sim e^{\pi\sqrt{q_e^2 q_m^2 - (q_e\cdot q_m)^2}}. \tag{2.9}$$

A formal derivation of this result is outlined in the Appendix, and makes use of the zeroes of the function Φ. These zeroes lead to poles in the integrand in (2.3). In our case the leading contribution to the integral comes from the pole at $\rho\sigma - \upsilon^2 + \upsilon = 0$(which is an appropriate $Sp(2, \mathbf{Z})$-transform of the above described 'tachyon pole'). After evaluating the residue at this pole one can perform the remaining integration by means of a saddle-point approximation, and this immediately reproduces the above formula for the asymptotic growth of $d(q_e, q_m)$. We will give an alternative explanation of this result at the end of the next section, after we have reinterpreted the degeneracy formula (2.3) in terms of strings living on a five-brane.

3. DERIVATION FROM THE QUANTUM FIVEBRANE

It is known that the heterotic string arises as a soliton in type II string theory that can be described as a five-brane wrapped around the internal $K3$ manifold [12]. $SL(2, \mathbf{Z})$ symmetry in the type II formulation is a consequence of T-duality, and thus string-string duality implies that both the electric and magnetic heterotic string can be considered as five-branes wrapped around $K3$. Based on the above the degeneracy formula, we will argue in the following that this correspondence can be naturally extended to all dyonic states.

Heterotic string from the five-brane

The five-brane we will consider arises as a soliton in the 10-dimensional type II effective string theory [13]. The long-wavelength fields on its world-volume are given by an $N = (2,0)$ supermultiplet consisting of a tensor with self-dual 3-index field strength T_3, 4 chiral fermions and 5 scalar fields. These fields represent the zero modes of the 10-dimensional massless fields on the five-brane background, which, via the standard adiabatic argument, are allowed to vary along the five-brane world-volume.

The relation of the five-brane with the heterotic string can be derived by dimensionally reducing the six-dimensional chiral field theory on the world-volume to two dimensions on $K3$. In this way the self-dual tensor field reduces to 19 left-moving and 3 right-moving scalars. Together with the additional five scalars and fermions, this gives exactly the world-sheet fields of the heterotic string. Moreover, the string momentum lattice $\Gamma^{20,4}$ corresponds to the cohomology lattice of $K3$. In particular, the fluxes of the self-dual tensor field T_3 are identified with the electric charges

$$q_e^A = \int_{S^1 \times \Sigma^A} T_3 \tag{3.1}$$

where S^1 is one of the cycles of T^2 and Σ^A is a basis of two-cycles in $K3$.

In our setup, we would in fact like to take this correspondence one step further, by replacing the standard world-sheet theory of the heterotic string by the 5+1-dimensional world-volume theory of the five-brane wrapped around $K3$. Concretely this means that in counting the states of this generalized heterotic string, we wish to take seriously the possible fluctuations of the five-brane in the internal $K3$. At first sight this would appear to imply a very drastic extension of the world-sheet degrees of freedom, but it turns out that for purely electrically charged BPS states nothing has changed: when one imposes the BPS conditions one automatically finds that the allowed fluctuations of the five-brane are the usual left-moving heterotic string oscillations [14]. We will outline this calculation in a moment, after we have explained our general set-up.

Idea of the computation

We would like to extend the above description to include dyonic states. As we will argue in the following, these dyonic states will naturally arise as 1/4 BPS configurations of the five-brane, provided we make one important new assumption. Namely, we will assume that the world volume description of the five-brane is not described by a field theory, but rather by an appropriate closed string theory. This world-volume string theory must satisfy the condition that the usual massless collective fields should correspond to the ground states of this string, in direct analogy with the standard relation between the space-time effective field theory and the ground states of the type II string. In this sense, we wish to think about the strings on the five-brane as essentially the original type II string, but restricted to the world-volume. The effective field theory of the string ground states is sufficient to describe the long-wavelength fluctuations of the five-brane, but at short distance we will also take into account the other string modes, including winding configurations and excitations.

Our aim is eventually to determine the degeneracy of BPS states. We will do this by considering the partition function of the 5+1-dimensional world-volume string theory with target space $K3 \times T^2$.[*] In a light-cone gauge in which the transversal space is taken to be the $K3$ manifold, the worldsheet theory on this string is a $N = (4,4)$ superconformal field theory with $SU(2)_L \times SU(2)_R$ current algebra [15, 14]. The eight supercharges on the world-sheet describe the unbroken part of the space-time supersymmetry. (The broken space-time supersymmetries give rise to zero-modes on the five-brane.)

The five-brane BPS partition function $\mathcal{Z}$ essentially counts the number of BPS states of this second quantized string theory. In computing this quantity, we will neglect string interactions. The justification for this should come from the fact that the BPS-condition in combination with the large number of supersymmetries severely restricts the possible interactions. This expectation will indeed find concrete support in our analysis, as it will turn out that for the relevant string partition sums on the target space $K3 \times T^2$, there are no higher loop correction to the free energy. Hence the fivebrane BPS partition sum can be computed by taking the exponent of the relevant one-loop free energy of the string on the world-volume, provided one also adds an appropriate zero-mode contribution

---

[*] This calculation is the direct analog of the conventional counting of string states by means of a conformal field theory partition function on a world-sheet of topology T^2.

$$\mathcal{Z} = \exp\Big[\mathcal{F}^{string}_{1-loop} + \mathcal{F}_{0-modes}\Big]. \tag{3.2}$$

The concrete form of this zero-mode contribution follows from considering the free energy of the fluxes of the low-energy massless field theory on the five-brane world-volume. The partition sum $\mathcal{Z}$ will depend in particular on the moduli that parametrize the shape and size of world-volume manifold.

To obtain the number of BPS states for given electric and magnetic charge, one finally needs to identify the corresponding contribution to the partition sum. In our formalism, this identification is automatically implemented via the integral over an appropriate set of world-volume moduli, which act as Lagrange multipliers that impose 'level-matching conditions' analogous to the $L_0 - \overline{L}_0 = 0$ condition in string theory. We will now explain and motivate this procedure in more detail for the 1/2 BPS states, in which case we should recover the before-mentioned classical correspondence with the electric and magnetic heterotic string.

Electric and magnetic fivebrane states

To count the states that respect half of the supersymmetries we have to restrict to the string ground states. This turns the 6-dimensional string into a topological string for which we can simply compute its one-loop free energy $\mathcal{F}$. The one-loop computation is well-known [16]. There are 24 ground states on $K3$, corresponding to the harmonic forms. We further add the momenta and winding numbers in the T^2 direction, which take value in the lattice $\Gamma^{2,2}$ parametrised by the moduli ρ, σ. Here ρ parametrize the size and B-field on the two-torus and σ parametrizes the complex structure. We then compute [16]

$$\begin{aligned}\mathcal{F} &= \tfrac{1}{2}\int \frac{d^2\tau}{\tau_2} \sum_{(p_L,p_R)\in\Gamma^{2,2}} e^{i\pi(\tau p_L^2 - \overline{\tau} p_R^2)} \cdot 24 \\ &= 24 \log\Big(\rho_2^{1/2}|\eta(\rho)|^2\Big) + 24\log\Big(\sigma_2^{1/2}|\eta(\sigma)|^2\Big) + cnst.\end{aligned} \tag{3.3}$$

Here the first term represents the contribution of the winding modes of the strings that couple to the parameter ρ and the second term the contribution of the momentum modes that couple to the parameter σ. The occurrence of these two terms reflects a $\mathbf{Z}_2$ symmetry, which is an $R \to 1/R$ symmetry applied to the world-volume of the five-brane. This transformation interchanges momentum and winding modes. So, even at the level of the string ground states, we encounter a pure stringy effect.

To obtain the 1/2 BPS partition function there are two more steps to be taken. First, we have to take only the holomorphic part of the above expression in ρ and σ, and secondly, we need to include the zero-mode contributions. Specifically, we must replace the factor σ_2 in $\mathcal{Z} = e^{\mathcal{F}}$ by the appropriate theta-function representing the summation of the fluxes over the lattice $\Gamma^{22,6}$, see [17]. Combined with the momentum contribution $1/\eta(\sigma)^{24}$ this then gives the left-moving partition function of the heterotic string with modular parameter σ. This is in accord with the expected one-to-one correspondence between 1/2 BPS states of the five-brane and those of the heterotic string.

To restore the $\mathbf{Z}_2$ symmetry between ρ and σ, we also need to replace the powers of ρ_2 by a lattice sum with a new set of charges. It seems natural to identify these new charges with the magnetic charges $q_m \in \Gamma^{22,6}$. This identification finds support from the heterotic-type II string duality, as follows. The electric-magnetic duality map that interchanges q_e and q_m acts on the type IIA side as a T-duality on the T^2. Since the five-brane we consider is wrapped around $K3$ and a one-cycle in T^2, this transformation will act non-trivially on its world-volume via an $R \to 1/R$ transformation on the S^1. As we have just discussed, this is a symmetry of the world-volume string theory that interchanges the world-volume parameter ρ and the parameter σ.

Hence, we conclude that by representing the world-volume degrees of freedom on the five-brane as strings, we are able to describe the electrically charged heterotic string and the magnetically charged heterotic string both in terms of one single fivebrane.

Dyonic fivebrane states

We now wish to consider five-brane states that carry both electric and magnetic charge. These dyonic states necessarily break 3/4 of the 16 supercharges. Concretely this means that the dyonic BPS five-branes carry more degrees of freedom than the purely electric or magnetic configurations.

As already noted, the 8 unbroken supercharges of the five-brane are realized on the world-sheet of the string as the 4 left-moving and 4 right-moving generators of an $N = (4,4)$ superconformal algebra. So to keep 4 unbroken charges we must restrict the string states to be in the ground state of either the right-moving or the left-moving sector. For definiteness, we will choose the right-moving one. The index that counts these string states is given by the elliptic genus of $K3$ [18], which is defined as the weighted trace in the RR-sector of the superconformal sigma-model

$$\chi_{\tau,z}(K3) = \mathrm{Tr}\,(-1)^{F_L+F_R} e^{2\pi i(\tau(L_0-\frac{c}{24})+zF_L)} \tag{3.4}$$

with $c = 6$ for $K3$. The fermion numbers F_L and F_R can be identified with the zero-modes of the left-moving and right-moving $U(1) \subset SU(2)$ current algebras.

As we restrict to right-moving ground states, F_R only takes the values $0, \pm 1$, while F_L can a priori take all integer values and gives us a conserved quantum number m. Since the corresponding $SU(2)$-symmetry is part of the space-time Lorentz symmetry, this conserved quantum number can be interpreted as the helicity in space-time, *cf* [19, 20]. In the following we will treat it on equal footing with the momentum and winding numbers of the string around T^2. This turns the Narain lattice $\Gamma^{2,2}$ of the two-torus into the lattice $\Gamma^{3,2}$. Moreover, it allows us to extend the set of world-brane parameters ρ, σ by another complex parameter υ that couples to the helicity quantum number m. Technically, υ has an interpretation as a Wilson loop on the world-volume that parametrizes the $SU(2)_L$ bundle over T^2. Together, ρ, σ and υ parametrize the lattice $\Gamma^{3,2}$, and transform in a vector representation of the $SL(2,\mathbf{Z})$ subgroup of the corresponding $SO(3,2;\mathbf{Z})$ T-duality group. (See Appendix.)

We now need to compute the following one-loop integral

$$\mathcal{F} = \frac{1}{2}\int \frac{d^2\tau}{\tau_2} \sum_{\substack{(p_L,p_R)\,\in\,\Gamma^{3,2} \\ n\,\in\,4\mathbf{Z}-\epsilon}} e^{i\pi(\tau p_L^2 - \bar\tau p_R^2)} \cdot c(n) e^{i\pi\tau n/2} \tag{3.5}$$

where $\epsilon = 0, 1$ depending on whether the helicity quantum number m is even or odd, and the coefficients $c(n)$ are defined by the expansion of the $K3$ elliptic genus as[†]

$$\chi_{\tau,z}(K3) = \sum_{h\geq 0,\, m\in\mathbf{Z}} c(4h - m^2) e^{2\pi i(h\tau + mz)} \tag{3.6}$$

It turns out that precisely the integral (3.5) has been computed recently by T. Kawai in the context of $N = 2$ heterotic string compactifications [21] along the lines of [22]. The answer obtained in [21] reads

$$\mathcal{F} = -\log\left[(\rho_2\sigma_2 - \upsilon_2^2)^{10}|\Phi(\rho,\sigma,\upsilon)|^2\right] + const. \tag{3.7}$$

where the holomorphic part has the following product representation

$$\Phi(\rho,\sigma,\upsilon) = e^{2\pi i(\rho+\sigma+\upsilon)} \prod_{(k,l,m)>0} \left(1 - e^{2\pi i(k\rho + l\sigma + m\upsilon)}\right)^{c(4kl-m^2)}. \tag{3.8}$$

Here $(k,l,m) > 0$ means that $k, l \geq 0$ and $m \in \mathbf{Z}$, $m < 0$ for $k = l = 0$. (Note further that $c(n) = 0$ for $n < -1$).

[†] Up to normalization this is the unique weak Jacobi form of index 1 and weight 0.

Now, following the same logic as for the 1/2 BPS states, we define the partition sum for the dyonic 1/4 BPS states by $\mathcal{Z} = e^{\mathcal{F}}$, where again we have to include separately the zero-mode contributions and restrict to the holomorphic sector. So the denominator of the partition function is given by $\Phi(\rho, \sigma, \upsilon)$ and represents the contribution from multi-string states that describe the fluctuations of the five-brane preserving the BPS condition. From the appearance of the factor $\rho_2\sigma_2 - \upsilon_2^2 = \det \operatorname{Im} \Omega$ it is clear that the zero-modes must be included by a factor

$$\exp\Big[i\pi(q_m^2\rho + q_e^2\sigma + 2q_e \cdot q_m\upsilon)\Big] \tag{3.9}$$

Combining all ingredients, and continuing the parallel with the heterotic string discussion of the introduction, we finally arrive at the announced result for the (bosonic minus fermionic) degeneracy formula

$$d(q_e, q_m) = \oint d\rho d\sigma d\upsilon \, \frac{e^{i\pi(q_m^2\rho + q_e^2\sigma + 2q_e\cdot q_m\upsilon)}}{\Phi(\rho, \sigma, \upsilon)}. \tag{3.10}$$

As was also noted in [21], it is natural to combine the moduli ρ, σ and υ into a matrix Ω as in (2.2). In the above form, the function $\Phi(\rho, \sigma, \upsilon)$ is manifestly $SL(2, \mathbf{Z})$ invariant, which suggests that it can be written as a modular form in Ω. Indeed, a remarkable identity proved by Gritsenko and Nikulin [23] expresses the product representation (3.8) in terms of the product of even genus 2 theta-functions as in formula (2.4). The role of the automorphic form Φ in string theory was first recognized in [24].

State counting

The above single five-brane partition function has a concrete interpretation as a trace of the multi-string quantum states of a target space topology $K3 \times S^1 \times \mathbf{R}$. On such a target space a string carries three additive quantum numbers: a winding number k, a momentum l along the S^1 and a helicity m. Together the three integers (k, l, m) form a vector on the lattice $\Gamma^{2,1}$ with length-squared $m^2 - 4kl$, which is invariant under the symmetry group

$$SO(2, 1, \mathbf{Z}) \cong SL(2, \mathbf{Z}) \tag{3.11}$$

We claim that this $SL(2, \mathbf{Z})$ group should be identified with the electric-magnetic duality group. Level-matching on the string tells us that the conformal dimension h in the $K3$ sigma model must be set equal to $h = kl$. If we look at the elliptic genus formula we see that the number of single string states with given quantum numbers (k, l, m) is $|c(4kl - m^2)|$, and thus duality invariant. Such a string state represents a space-time boson or fermion depending on the sign of $c(4kl - m^2)$. (The sign of $c(n)$ is $+1$ if $n \equiv 0 \pmod 4$ and -1 if $n \equiv -1 \pmod 4$, with the exception of $c(-1) = 2$.)

A state in the multi-string Hilbert space is characterized by a set of occupation numbers $N^I_{k,l,m}$ with $I = 1, \ldots, |c(4kl - m^2)|$. Hence, the degeneracy $d(q_e, q_m)$ counts the number of independent multi-string states $|\{N^I_{k,l,m}\}\rangle$ that satisfy the appropriate level-matching conditions implemented via the integral over the moduli σ, ρ and υ. These take the form

$$\begin{aligned} \tfrac{1}{2}q_m^2 + \sum_{k,l,m,I} kN^I_{k,l,m} &= 1 \\ \tfrac{1}{2}q_e^2 + \sum_{k,l,m,I} lN^I_{k,l,m} &= 1 \\ q_e \cdot q_m + \sum_{k,l,m,I} mN^I_{k,l,m} &= 1 \end{aligned} \tag{3.12}$$

The first two relations are obvious generalisations of the level-matching conditions of the heterotic string. In the case of only electric or magnetic charges they indeed reproduce the usual counting. The third line relates the sum of the helicities of the individual strings to the quantity $q_e \cdot q_m$. A suggestive interpretation of this constraint is that $q_e \cdot q_m$ represents the contribution to the helicity as carried by the external electric-magnetic field of the dyon, when considered as made up from a separate electric and magnetic 'point charge'. According to this interpretation, the third constraint in (3.13) essentially puts the total helicity of the dyonic state equal to zero.

4. COMPARISON WITH D-BRANE COUNTING.

Here we would like to indicate how this representation of the BPS states is related to the more conventional construction in terms of D-branes [25, 26, 7, 9]. As we will argue, both the string representation as well as the level matching relations have a natural interpretation from this point of view.

Strings and D-brane intersections

Let us consider a five-brane wrapped around the $K3$ compactification manifold, so that its world-volume has the space-like topology of $K3 \times S^1$. A pure five-brane only carries NS-NS charge. More general BPS-configurations that also carry R-R charge can be obtained by forming bound states with D-branes. For our purposes the most convenient description of these bound states is given in the context of type II B string theory. In this case we need to include 1-branes, 3-branes and 5-branes. When more of these D-branes are taken to lie within the five-brane world volume, they will in general have intersections [26, 25]. The collection of BPS states that corresponds to a given charge configuration are obtained by quantizing the appropriate degrees of freedom of these intersections [7].

Since $K3$ has only non-trivial homology cycles of even dimension, on the given topology of the five-brane of $K3 \times S^1$, all the type IIB D-branes will be wrapped around a corresponding element of $H_*(K3)$ times the S^1. Hence to a given D-brane configuration we can associate a vector

$$q \in H_*(K3, \mathbf{Z}) \tag{4.1}$$

in the integral homology of $K3$. This vector becomes identified with the charge vector

$$q \in \Gamma^{20,4} \tag{4.2}$$

via the isomorphism of $H_*(K3, \mathbf{Z}) \cong \Gamma^{20,4}$, in which the norm on $\Gamma^{20,4}$ is identified with the intersection form. A pair of 3-branes will typically intersect along a 1-dimensional string wound one or more times around the S^1 direction, and similarly the 1-brane intersects with the 5-brane along a string around the S^1.

These strings formed by the D-brane intersections encode the internal degrees of freedom of the five-brane in a completely similar fashion as the string introduced in section 3. In particular, its transversal oscillations are also described by an appropriate $N = 4$ superconformal field theory with target space $K3$, where the BPS-restriction again only allows for either left- or right-moving string oscillations. The spectrum of these quantized 'intersection-strings' will in particular include the excitations of the low-energy effective field theory on the five-brane world volume. The number of excited string states is again counted by means of the $K3$ eliptic genus.

It is evident from the above description that the total string winding number is given by the intersection pairing of the homology class (4.1) on $K3$ defined by the full D-brane configuration. Hence we see that the D-brane representation gives a clear geometrical interpretation of the matching condition

$$\tfrac{1}{2}q_m^2 + \sum_{k,l,m,I} k N^I_{k,l,m} = 1. \tag{4.3}$$

Here we identified the vector in (4.1) with the magnetic charge q_m.

In addition to the string winding number, the string oscillations will also contribute to the total momentum along the S^1. This total momentum is measured from the perspective of the five-brane world-volume, and is therefore related but not equal to the internal KK momentum. The relation between the two quantities is determined via a level matching condition equating this total string momentum to a zero-mode contribution. As before, the form of the zero-mode contribution is determined by means of the free field theory of the massless five-brane modes, and thus depends quadratically on the fluxes and KK-momenta that gives rise to the electric charge q_e. This gives the second condition

$$\tfrac{1}{2}q_e^2 + \sum_{k,l,m,I} l N^I_{k,l,m} = 1. \tag{4.4}$$

as the constraint of translation invariance on the world-volume along the S^1 direction.

Finally, since the D-brane intersection strings carry a non-zero space-time helicity m, we can also introduce a third level matching condition that equates the sum of the individual string contributions to the total helicity quantum number.

T-duality

We can learn more about the dyonic five-brane configuration by considering the effect of the T-duality transformation that interchanges the electric and magnetic charges. Concretely, let us consider the configuration of a dyonic five-brane wound along the A-cycle of the internal two-torus T^2, and act on it with the simultaneous $R \to 1/R$ transformation on both the A- and B-cycle. This map leaves the five-brane location invariant. However, when applied to the intersection strings within the five-brane, it changes their winding direction to the B-cycle, while in addition it interchanges the winding and momentum quantum numbers. It would appear therefore that the new intersection strings lie perpendicular to the five-brane.

An alternative conclusion, however, is that a dyonic five-brane should rather be thought of a pair of perpendicular five-branes, one around the A-cycle and one around the B-cycle. Both five-branes contain a collection of D-brane intersection strings, which however are T-duals of each other in that the momentum quantum numbers of the A-cycle strings are identified with the winding quantum numbers of the B-cycle strings. In this picture, the total winding number of the intersection strings around, say, the A-cycle is measured by q_m^2, while the total string winding number around the B-cycle is q_e^2.

A generating function

It has been argued that the above multiple D-brane description leads to a two-dimensional sigma model with target space the sum of symmetric products of $K3$ manifolds [7]. The 1/2 BPS states are in one-to-one correspondence with ground states of this model and correspond to harmonic forms. The number of these ground states follows from a direct application of the so-called orbifold formula, and is summarized in terms of the generating function (see [19])

$$\sum_{n} e^{2\pi i n\sigma} \dim H^{*}\left(\frac{(K3)^{n}}{S_{n}}\right) = \prod_{l>0} \frac{1}{(1-e^{2\pi i l\sigma})^{24}} \tag{4.5}$$

Our results now suggest a concrete generalization of this counting formula to the 1/4 BPS states, in which case one has to include the transversal string oscillations.

As was first suggested in [7], the counting of these states is naturally related to the elliptic genus of the above-mentioned sigma model. Based on our degeneracy formula, we conjecture that these elliptic genera have a generating function [*]

$$\sum_{n} e^{2\pi i n\sigma} \chi_{\rho,\upsilon}\left(\frac{(K3)^{n}}{S_{n}}\right) = \prod_{k\geq 0, l>0, m\in \mathbf{Z}} \frac{1}{(1-e^{2\pi i(k\rho+l\sigma+m\upsilon)})^{c(4kl-m^{2})}} \tag{4.6}$$

This conjectured identity suggests the following physical interpretation. The terms on the left-hand side count the oscillations of a single long string on the multiple product of the transversal target space $K3$. The right-hand side of our formula (4.6) on the other hand gives the usual second quantized representation of a multiple string state on $K3$ in terms of a string field theory Fock space. The equality of both sides, if indeed true, indicates that the single long string state on the symmetric product of $K3$ provides an exact first quantized representation of the appropriately symmetrized multiple string state on $K3$. Loosely speaking, we equate the elliptic genus of the 'second-quantized manifold' to the second-quantized elliptic genus of a single copy of the manifold. Based on this intuition, we expect that in the above relation the manifold $K3$ can be replaced by an arbitrary manifold X, if one uses the expansion coefficients $c(n)$ of the elliptic genus of X.

We note further that the above formula satisfies a number of highly non-trivial consistency checks. First, the $n = 1$ term gives indeed the answer (3.6) for the $K3$ elliptic genus. Secondly, for $\upsilon = 0$ one recovers the ground state expression (4.5). Thirdly, the expression on the right-hand side of equation (4.6) has the required automorphic properties. Precisely because of that reason this particular expression has been considered before in the literature [23]. It has the important property that, when expanded in powers of $e^{2\pi i\sigma}$ (the so-called Fourier-Jacobi expansion) the n-th coefficient is a weak Jacobi form of weight zero and index n for the modular group $SL(2,\mathbf{Z})$ that acts on the modulus ρ and υ. (See [27] for the theory of Jacobi forms.) The expansion coefficients are therefore indeed candidates for elliptic genera of Calabi-Yau manifolds of (complex) dimension $2n$, which matches the expansion on the left-hand side. Note that the space of weak Jacobi forms of weight zero and fixed index is finite-dimensional. Given the fact that the Euler characters match, this gives extra support to our conjecture.

---

[*] Here in the product we removed, relative to (3.8), the contribution of the states with $l = 0$. The resulting expression is exactly the contribution of the non-degenerate orbits in the fundamental domain integral [21].

Finally, we note that the asymptotic growth of the right-hand side, as described in section 2, gives via the above equality a prediction for the number of single string states on $(K3)^n/S_n$ at large oscillation level. This prediction can be verified by standard techniques.

5. CONCLUDING REMARKS

Let us finally mention some applications and possible generalizations of our result.

Five-dimensional degeneracies

One of the consequences of our analysis is a concrete refinement of the results of Strominger and Vafa [7] on the counting of microscopic black hole states in five dimensions. The formula (4.6) gives an explicit counting formula of these states and in particular allows one to analyze the asymptotic growth in more directions by separately controling the total string winding, momentum and helicity number. In five dimensions the level matching conditions equate these quantum numbers to respectively q^2, the charge Q_H of [7] and the five-dimensional spin J. Explicitely, the degeneracy of these states is given by

$$d(q, Q_H, J) = \oint d\rho d\sigma d\upsilon \, \frac{e^{i\pi(q^2\rho + Q_H\sigma + 2J\upsilon)}}{\Phi'(\rho, \sigma, \upsilon)} \tag{5.1}$$

with Φ' the denominator on the right-hand side of equation (4.6). The asymptotics of Φ' are similar to those of Φ as analyzed in the Appendix.

Generalizations

It would be interesting to extend our results to more general four-dimensional black hole states, by allowing for the possibility of non-zero angular momentum [20] or states away from extremality.

Spinning extremal black holes are in fact most naturally included in our description, although for this we would need to give up manifest $SL(2, \mathbf{Z})$-duality. The most general spinning dyonic black hole geometry is characterized by *four* independent macroscopic quantities (namely the angular momentum J, the length-squared of the charges q_e^2, q_m^2, and $q_e \cdot q_m$). Our most general states, on the other hand, carry at most three macroscopic quantum numbers. In our presentation we eventually arrived at a duality symmetric expression by equating the total 'internal' helicity to the 'external' helicity $q_e \cdot q_m$. However, we can in principle omit this constraint and simply identify the total internal helicity with the macroscopic external angular momentum. This will give a degeneracy formula consistent with the predicted entropy for spinning black holes [20] in the case that $q_e \cdot q_m = 0$.

Another important generalization is to include near extremal black hole states [28]. Within our framework, an obvious step is to relax the BPS condition on the individual strings, and allow for arbitrary left- and right-moving oscillations. We expect that a naive extension of our formalism in this regime would indeed produce an asymptotic number of states in accordance with the four-dimensional macroscopic entropy proportional to the mass squared. This is in essence a consequence of the fact that the mass-shell relation $L_0 =(\mathrm{mass})^2$ is applied twice in our formalism: once on the string on the world-volume, and once on the five-brane in space-time. By a similar argument we also expect that the energy gap will also behave in accordance with the thermodynamic prediction. We should stress however that the resulting description will very likely be of limited validity, since on various crucial points we have made use of the BPS condition. In particular it is not at all clear that the (effective) string theory on the world-volume can still be treated as an (approximately) free string theory.

Algebraic structure

Our results suggest that there exists an interesting algebraic structure behind the $N = 4$ dyon spectrum, related to a generalized Kac-Moody (GKM) algebra. The particular (super)algebra that underlies our degeneracy formula is a so-called automorphic correction [23] of a hyperbolic Kac-Moody algebra with three simple roots with Cartan matrix

$$\begin{pmatrix} 2 & -2 & -2 \\ -2 & 2 & -2 \\ -2 & -2 & 2 \end{pmatrix} \tag{5.2}$$

(Note that the corresponding GKM has an infinite number of imaginary simple roots, roughly corresponding to the single string BPS states.) The fact that the denominator of this GKM appears (in the denominator) of our partition function, suggests that perhaps the full characters, including the numerator, also play a role in the BPS spectrum. Perhaps it is possible to realize the spectrum generating algebra of the $N = 4$ dyon BPS states in terms of this symmetry algebra (cf. [29]). Generalized Kac-Moody algebras also recently appeared in the $N = 2$ context [22], where it was suggested that the heterotic string vertex operators representing the perturbative BPS states form a GKM.

ACKNOWLEDGEMENTS

We would like to thank S. Bais, R. Borcherds, S. Ferrara, G. v.d. Geer, C. Hofman, E. Kiritsis, C. Kounnas, G. Moore, D. Neumann, L. Susskind, C. Vafa and E. Witten for helpful discussions. This research is partly supported by a Pionier Fellowship of NWO, a Fellowship of the Royal Dutch Academy of Sciences (K.N.A.W.), the Packard Foundation and the A.P. Sloan Foundation.

APPENDIX

The asymptotic behavior of the degeneracy formula (2.3) are related to the zeroes of the automorphic form Φ in the integrand. These zeroes are well understood using either standard facts on genus two Riemann surfaces, the results of Borcherds [10] or the representation in terms of a string one-loop amplitude [22, 21] as in (3.5). From the latter point of view we understand that $\Phi(\rho,\sigma,\upsilon)$ acquires a zero or pole whenever the moduli ρ, σ and υ that parametrize the Narain lattice $\Gamma^{3,2}$ allow chiral vertex operators, *i.e.* elements $p=(p_L,p_R)\in\Gamma^{3,2}$ with $p_R=0$. In terms of the integer lattice vector $p=(k,l,m,a,c)\in\Gamma^{3,2}\cong\mathbf{Z}^5$ we have [21]

$$p_R^2=\frac{1}{2Y}|k\rho+l\sigma+m\upsilon+a(\rho\sigma-\upsilon^2)+c|^2 \tag{A.1}$$

with $Y=\rho_2\sigma_2-\upsilon_2^2$, and

$$p^2=\tfrac{1}{2}m^2-2kl+2ac. \tag{A.2}$$

If we use the vector notation $y=(\rho,\sigma,2\upsilon)$, with $-\frac{1}{2}y^2=\rho\sigma-\upsilon^2$, we see that the zeroes and poles occur at the 'rational quadratic divisors' of [10]

$$-\tfrac{1}{2}ay^2+b\cdot y+c=0 \tag{A.3}$$

where $b=(k,l,m)\in\Gamma^{2,1}$ with norm-squared $b^2=-2kl+\frac{1}{2}m^2$ and $a,c\in\mathbf{Z}$. We can write equation (A.3) also as $p\cdot v=0$, where we have introduced the 5-dimensional null vector $v=(y;-\frac{1}{2}y^2,1)$. Of course $|p\cdot v|^2\sim p_R^2$. In terms of the coefficients $c(n)$, the order of the zero of $\Phi(y)$ at such a divisor is $\sum_{n>0}c(-2n^2p^2)$ with $p^2>0$. In our case the only possibility is $c(-1)=2$, so we necessarily have $p^2=1/2$ and the partition function $1/\Phi(y)$ has a second order pole at this divisor.

In the theory of genus two curves or abelian surfaces, the equation (A.3) is well-known to describe (a component of) the so-called Humbert surface H_Δ with discriminant $\Delta=2p^2$ [30]. In our case we are interested in the surface H_1. This is indeed known to be the zero-locus of the product of all even theta-functions.

The 'tachyon pole' at $\upsilon = 0$ occurs for $b = (0,0,1)$ and $a = c = 0$, which describes the factorization on a product of two elliptic curves. All the other poles are transformations of this one by $Sp(2, \mathbf{Z}) \cong SO(3,2,\mathbf{Z})$. This group acts linear on the vectors v and p, which is another way to see that the divisors are of the form (A.3). The leading contribution in the Fourier coefficient

$$D(q) = \int d^3y \, \frac{e^{2\pi i q \cdot y}}{\Phi(y)} \tag{A.4}$$

for large q from the divisor (A.3) is of the form $D(q) \sim e^{2\pi\mu|q|}$ with $\mu^2 = (b^2 + 2ac)/a^2 = 1/2a^2$. So the dominant term will be given by $a = 1$. This gives indeed the black-hole entropy (2.9). The other coefficients are then constrainted by $4c - 4kl + m^2 = 1$, so in particular m is odd, and k, l can be arbitrary. However, since we have the freedom to shift the variables ρ, σ, υ by integers, we can put $k, l, c = 0$ and $m = 1$. The divisor is then given by $\rho\sigma - \upsilon^2 + \upsilon = 0$.

Let us finally make a few remarks about the geometrical interpretation in terms of a genus two surface. Let A_1, A_2, B_1, B_2 be a basis for the homology group $H_1(\Sigma)$ of a genus two Riemann surface Σ. Now consider the quotient of $\bigwedge^2 H_1$ by the relation

$$A_1 \wedge B_1 + A_2 \wedge B_2 = 0 \tag{A.5}$$

and call the resulting 5-dimensional lattice $L \cong \Gamma^{3,2}$. It naturally carries a signature $(3,2)$ bilinear form, coming from the intersection form on $\bigwedge^2 H_1$. This makes explicit the isomorphism $Sp(2,\mathbf{Z}) \cong SO(3,2,\mathbf{Z})$. As a basis of L we choose

$$B_1 \wedge A_2,\ B_2 \wedge A_1,\ A_1 \wedge B_1 = B_2 \wedge A_2,\ B_1 \wedge B_2,\ A_1 \wedge A_2. \tag{A.6}$$

There is a natural map $L \to \mathbf{C}$ given in terms of the periods of the abelian differentials ω_1, ω_2

$$C \wedge D \to \oint_C \omega_1 \oint_D \omega_2 - \oint_D \omega_1 \oint_C \omega_2 \tag{A.7}$$

Under this map our basis (A.6) gives the 'bi-periods'

$$\rho,\ \sigma,\ \upsilon,\ \rho\sigma - \upsilon^2,\ 1. \tag{A.8}$$

It is clear that symplectic transformations will act linear on these variables. To any basis element of L we can associate a zero-homology cycle in the homotopy group $\pi_1(\Sigma)$ by

$$C \wedge D \to [C, D] = CDC^{-1}D^{-1} \tag{A.9}$$

Note that $[A_1, B_1][A_2, B_2] \sim 1$ so that this map is well-defined. It associates a bi-period to a vanishing cycle. Indeed, in this way the divisor $\upsilon = 0$ describes the usual pinching along the dividing cycle $[A_1, B_1] = [B_2, A_2]$. Other divisors can be similarly interpreted.

References

[1] A. Font, L.E. Ibanez, D. Lust, and F. Quevedo, "Strong-Weak Coupling Duality and Nonperturbative Effects in String Theory," Phys.Lett. **B249** (1990) 35-43.

[2] A. Sen, "Strong-Weak Coupling Duality in Four Dimensional String Theory," Int.J.Mod.Phys. **A9** (1994) 3707-3750, hep-th/9402002; "Dyon-Monopole Bound States, Self-Dual Harmonic Forms on the Multi-Monopole Moduli Space, and $SL(2,Z)$ Invariance in String Theory, Phys.Lett. **B329** (1994) 217-221, hep-th/9402032.

[3] J.P. Gauntlett and J.A. Harvey, "S-Duality and the Spectrum of Magnetic Monopoles in Heterotic String Theory," hep-th/9407111.

[4] J. Bekenstein, Phys. Rev. **D7** (1973) 2333; Phys. Rev. **D9** (1974) 3292.
S. W. Hawking, Phys. Rev. **D13** (1976) 191.

[5] M. Cvetic and D. Youm, "Dyonic BPS Saturated Black Holes in Heterotic String on a Six Torus," Phys.Rev. **D53** (1996) 584-588; hep-th/9507090. M. Cvetic and A. Tseytlin, "Solitonic Strings and BPS Saturated Dyonic Black Holes," hep-th/9512031.

[6] J. Polchinski, " Dirichlet-Branes and Ramond-Ramond Charges," hep-th/9510017; J. Polchinski, S. Chaudhuri, and C. Johnson, "Notes on D-Branes," hep-th/9602052.

[7] A. Strominger, C. Vafa, "Microscopic Origin of the Bekenstein-Hawking Entropy," hep-th/9601029.

[8] F. Larsen and F. Wilczek, "Internal Structure of Black Holes," hep-th/9511064; M. Cvetic and A. Tseytlin, "Solitonic Strings and BPS Saturated Dyonic Black Holes," hep-th/9512031; G. Horowitz and A. Strominger, "Counting States of Near-Extremal Black Holes," hep-th/9602051; C. Callan and J. Maldacena, "D-brane Approach to Black Hole Quantum Mechanics," hep-th/9602043.

[9] C. Johnson, R. Khuri and R. Myers, "Entropy of 4D Extremal Black Holes," hep-th/9603061; J. Maldacena and A. Strominger, "Statistical Entropy of Four-Dimensional Extremal Black Holes," hep-th/9603060.

[10] R. E. Borcherds, "Automorphic Forms on $O_{s+2,2}(R)$ and Infinite Products," Invent. Math. **120** (1995) 161.

[11] G. Moore, "Modular forms and two-loop string physics," Phys.Lett.**176B** (1986) 369; A. Belavin, V.Knizhnik, A. Morozov and A. Perelemov, Phys. Lett. **B 177** (1986) 324.

[12] J. A. Harvey and A. Strominger, "The Heterotic String is a Soliton," Nucl. Phys. **B 449** (1995) 535-552; Nucl. Phys. B458 (1996) 456-473, hep-th/9504047.

[13] M.J. Duff, R.R. Khuri, and J.X. Lu, "String Solitons," Phys.Rept. **259** (1995) 213-326, hep-th/9412184. C. Callan, J. Harvey, and A. Strominger, Nucl.Phys. **B367** (1991) 60.

[14] R. Dijkgraaf, E. Verlinde and H. Verlinde, "BPS Quantization of the Five-Brane," hep-th/9604055.

[15] R. Dijkgraaf, E. Verlinde and H. Verlinde, "BPS Spectrum of the Five-Brane and Black Hole Entropy," hep-th/9603126.

[16] L. Dixon, V. Kaplunovsky, and J. Louis, "Moduli-dependence of string loop corrections to gauge coupling constants," Nucl. Phys. **B307** (1988) 145.

[17] E. Verlinde, "Global Aspects of Electric-Magnetic Duality," Nucl.Phys. **B455** (1995) 211-228, hep-th/9506011

[18] T. Kawai, Y. Yamada and S.-K. Yang, "Elliptic Genera and N=2 Superconformal Field Theory," Nucl. Phys. **B414** (1994) 191-212. T. Eguchi, H. Ooguri, A. Taormina, S.-K. Yang, "Superconformal Algebras and String Compactification on Manifolds with $SU(N)$ Holonomy," Nucl.Phys. **B315** (1989) 193.

[19] C. Vafa and E. Witten, "A Strong Coupling Test of S-Duality," Nucl.Phys. **B431** (1994) 3-77, hep-th/9408074.

[20] J.C. Breckenridge, R.C. Myers, A.W. Peet and C. Vafa, "D–branes and Spinning Black Holes," hep-th/9602065; J.C. Breckenridge, D.A. Lowe, R.C. Myers, A.W. Peet, A. Strominger and C. Vafa, "Macroscopic and Microscopic Entropy of Near-Extremal Spinning Black Holes," hep-th/9603078.

[21] T. Kawai, "$N = 2$ Heterotic String Threshold Correction, $K3$ Surface and Generalized Kac-Moody Superalgebra," hep-th/9512046.

[22] J. Harvey and G. Moore, "Algebras, BPS States, and Strings," Nucl.Phys.**B 463** (1996) 315-368

[23] V.A. Gritsenko and V.V. Nikulin, "Siegel Automorphic Form Corrections of Some Lorentzian Kac-Moody Algebras," alg-geom/9504006.

[24] P. Mayr and S. Stieberger, Phys. Lett. **B 355** (1995) 107, hep-th/9504129.

[25] C. Vafa,"Gas of D-Branes and Hagedorn Density of BPS States," hep-th/9511026, "Instantons on D-branes," hep-th/9512078; M. Bershadsky, V. Sadov, and C. Vafa, "D-Branes and Topological Field Theories," hep-th/9511222.

[26] A. Sen, "A Note on Marginally Stable Bound States in Type II String Theory" hep-th/9510229; "U Duality and Intersecting D-Branes," hep-th/9511026; "T-Duality of p-Branes," hep-th/9512062.

[27] M. Eichler and D. Zagier, "The Theory of Jacobi Forms" (Birkhäuser, 1985).

[28] G. Horowitz and A. Strominger, “Counting States of Near-Extremal Black Holes,” hep-th/9602051; I. R. Klebanov, A. A. Tseytlin, “Entropy of Near-Extremal Black p-branes,” hep-th/9604089; J. M. Maldacena, “Statistical Entropy of Near Extremal Five-branes,” hep-th/9605016.

[29] A. Giveon and M. Porrati, “Duality Invariant String Algebra and $D = 4$ Effective Actions,” Nucl.Phys. **B355** (1991) 422-454.

[30] G. van der Geer, “Hilbert Modular Surfaces” (Springer, 1980).

E_8 (mathematics)

In mathematics, **E_8** is any of several closely related exceptional simple Lie groups, linear algebraic groups or Lie algebras of dimension 248; the same notation is used for the corresponding root lattice, which has rank 8. The designation E_8 comes from the Cartan–Killing classification of the complex simple Lie algebras, which fall into four infinite series labeled A_n, B_n, C_n, D_n, and five exceptional cases labeled E_6, E_7, E_8, F_4, and G_2. The E_8 algebra is the largest and most complicated of these exceptional cases.

Contents

- Basic description
- Real and complex forms
- E_8 as an algebraic group
- Constructions
- Geometry
- E_8 root system
  - Construction
  - Dynkin diagram
  - Cartan matrix
  - Simple roots
  - Weyl group
  - E_8 root lattice
  - Simple subalgebras of E_8
- Chevalley groups of type E_8
- Subgroups
- Applications
- History
- See also
- Notes
- References
- External links

Basic description

The Lie group E_8 has dimension 248. Its rank, which is the dimension of its maximal torus, is eight.

Therefore, the vectors of the root system are in eight-dimensional Euclidean space: they are described explicitly later in this article. The Weyl group of E_8, which is the group of symmetries of the maximal torus which are induced by conjugations in the whole group, has order $2^{14}\,3^5\,5^2\,7 = 696\,729\,600$.

The compact group E_8 is unique among simple compact Lie groups in that its non-trivial representation of smallest dimension is the adjoint representation (of dimension 248) acting on the Lie algebra E_8 itself; it is also the unique one which has the following four properties: trivial center, compact, simply connected, and simply laced (all roots have the same length).

There is a Lie algebra E_k for every integer $k \geq 3$. The largest value of k for which E_k is finite-dimensional is $k=8$, that is, E_k is infinite-dimensional for any $k > 8$.

Real and complex forms

There is a unique complex Lie algebra of type E_8, corresponding to a complex group of complex dimension 248. The complex Lie group E_8 of complex dimension 248 can be considered as a simple real Lie group of real dimension 496. This is simply connected, has maximal compact subgroup the compact form (see below) of E_8, and has an outer automorphism group of order 2 generated by complex conjugation.

As well as the complex Lie group of type E_8, there are three real forms of the Lie algebra, three real forms of the group with trivial center (two of which have non-algebraic double covers, giving two further real forms), all of real dimension 248, as follows:

- The compact form (which is usually the one meant if no other information is given), which is simply connected and has trivial outer automorphism group.
- The split form, EVIII (or $E_{8(8)}$), which has maximal compact subgroup Spin(16)/(**Z**/2**Z**), fundamental group of order 2 (implying that it has a double cover, which is a simply connected Lie real group but is not algebraic, see below) and has trivial outer automorphism group.
- EIX (or $E_{8(-24)}$), which has maximal compact subgroup $E_7 \times SU(2)/(-1,-1)$, fundamental group of order 2 (again implying a double cover, which is not algebraic) and has trivial outer automorphism group.

For a complete list of real forms of simple Lie algebras, see the list of simple Lie groups.

E_8 as an algebraic group

By means of a Chevalley basis for the Lie algebra, one can define E_8 as a linear algebraic group over the integers and, consequently, over any commutative ring and in particular over any field: this defines the so-called split (sometimes also known as “untwisted”) form of E_8. Over an algebraically closed field, this is the only form; however, over other fields, there are often many other forms, or “twists” of E_8, which are classified in the general framework of Galois cohomology (over a perfect field k) by the set $H^1(k,\mathrm{Aut}(E_8))$ which, because the Dynkin diagram of E_8 (see below) has no automorphisms, coincides with $H^1(k,E_8)$.[1]

Over **R**, the real connected component of the identity of these algebraically twisted forms of E_8 coincide with the three real Lie groups mentioned above, but with a subtlety concerning the fundamental group: all forms of E_8 are simply connected in the sense of algebraic geometry, meaning that they admit no non-trivial algebraic coverings; the non-compact and simply connected real Lie group forms of E_8 are therefore not algebraic and admit no faithful finite-dimensional representations.

Over finite fields, the Lang–Steinberg theorem implies that $H^1(k,E_8)=0$, meaning that E_8 has no twisted forms: see below.

The characters of finite dimensional representations of the real and complex Lie algebras and Lie groups are all given by the Weyl character formula. The dimensions of the smallest irreducible representations are (sequence A121732 in the OEIS):

> 1, 248, 3875, 27000, 30380, 147250, 779247, 1763125, 2450240, 4096000, 4881384, 6696000, 26411008, 70680000, 76271625, 79143000, 146325270, 203205000, 281545875, 301694976, 344452500, 820260000, 1094951000, 2172667860, 2275896000, 2642777280, 2903770000, 3929713760, 4076399250, 4825673125, 6899079264, 8634368000 (twice), 12692520960…

The 248-dimensional representation is the adjoint representation. There are two non-isomorphic irreducible representations of dimension 8634368000 (it is not unique; however, the next integer with this property is 175898504162692612600853299200000 (sequence A181746 in the OEIS)). The fundamental representations are those with dimensions 3875, 6696000, 6899079264, 146325270, 2450240, 30380, 248 and 147250 (corresponding to the eight nodes in the Dynkin diagram in the order chosen for the Cartan matrix below, i.e., the nodes are read in the seven-node chain first, with the last node being connected to the third).

The coefficients of the character formulas for infinite dimensional irreducible representations of E_8 depend on some large square matrices consisting of polynomials, the Lusztig–Vogan polynomials, an analogue of Kazhdan–Lusztig polynomials introduced for reductive groups in general by George Lusztig and David Kazhdan (1983). The values at 1 of the Lusztig–Vogan polynomials give the coefficients of the matrices relating the standard representations (whose characters are easy to describe) with the irreducible representations.

These matrices were computed after four years of collaboration by a group of 18 mathematicians and computer scientists, led by Jeffrey Adams, with much of the programming done by Fokko du Cloux. The most difficult case (for exceptional groups) is the split real form of E_8 (see above), where the largest matrix is of size 453060×453060. The Lusztig–Vogan polynomials for all other exceptional simple groups have been known for some time; the calculation for the split form of E_8 is far longer than any other case. The announcement of the result in March 2007 received extraordinary attention from the media (see the external links), to the surprise of the mathematicians working on it.

The representations of the E_8 groups over finite fields are given by Deligne–Lusztig theory.

Constructions

One can construct the (compact form of the) E_8 group as the automorphism group of the corresponding $\mathbf{e}_8$ Lie algebra. This algebra has a 120-dimensional subalgebra **so**(16) generated by J_{ij} as well as 128 new generators Q_a that transform as a Weyl–Majorana spinor of **spin**(16). These statements determine the commutators

$$[J_{ij}, J_{k\ell}] = \delta_{jk} J_{i\ell} - \delta_{j\ell} J_{ik} - \delta_{ik} J_{j\ell} + \delta_{i\ell} J_{jk}$$

as well as

$$[J_{ij}, Q_a] = \frac{1}{4}(\gamma_i\gamma_j - \gamma_j\gamma_i)_{ab} Q_b,$$

while the remaining commutators (not anticommutators!) between the spinor generators are defined as

$$[Q_a, Q_b] = \gamma^{[i}_{ac}\gamma^{j]}_{cb} J_{ij}.$$

It is then possible to check that the Jacobi identity is satisfied.

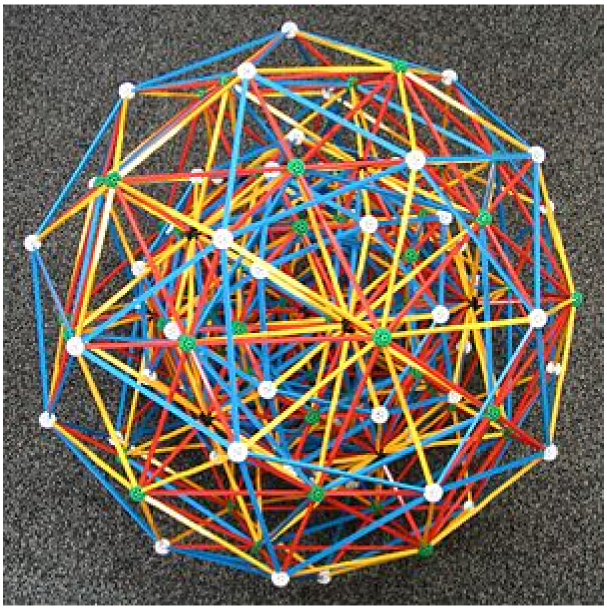
Geometry

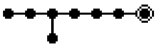
The compact real form of E_8 is the isometry group of the 128-dimensional exceptional compact Riemannian symmetric space EVIII (in Cartan's classification). It is known informally as the "octooctonionic projective plane" because it can be built using an algebra that is the tensor product of the octonions with themselves, and is also known as a Rosenfeld projective plane, though it does not obey the usual axioms of a projective plane. This can be seen systematically using a construction known as the *magic square*, due to Hans Freudenthal and Jacques Tits (Landsberg & Manivel 2001).

E_8 root system

A root system of rank *r* is a particular finite configuration of vectors, called *roots*, which span an *r*-dimensional Euclidean space and satisfy certain geometrical properties. In particular, the root system must be invariant under reflection through the hyperplane perpendicular to any root.

The $\mathbf{E_8}$ **root system** is a rank 8 root system containing 240 root vectors spanning $\mathbf{R}^8$. It is irreducible in the sense that it cannot be built from root systems of smaller rank. All the root vectors in E_8 have the same length. It is convenient for a number of purposes to normalize them to have length $\sqrt{2}$. These 240 vectors are the vertices of a semi-regular polytope discovered by Thorold Gosset in 1900, sometimes known as the 4_{21} polytope.



Zome model of the E_8 root system, projected into three-space, and represented by the vertices of the 4_{21} polytope, 

Construction

In the so-called *even coordinate system*, E_8 is given as the set of all vectors in $\mathbf{R}^8$ with length squared equal to 2 such that coordinates are either all integers or all half-integers and the sum of the coordinates is even.

Explicitly, there are 112 roots with integer entries obtained from

$$(\pm 1, \pm 1, 0, 0, 0, 0, 0, 0)$$

by taking an arbitrary combination of signs and an arbitrary permutation of coordinates, and 128 roots with half-integer entries obtained from

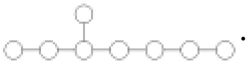
$$\left(\pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}, \pm\tfrac{1}{2}\right)$$

by taking an even number of minus signs (or, equivalently, requiring that the sum of all the eight coordinates be even). There are 240 roots in all.

The 112 roots with integer entries form a D_8 root system. The E_8 root system also contains a copy of A_8 (which has 72 roots) as well as E_6 and E_7 (in fact, the latter two are usually *defined* as subsets of E_8).

In the *odd coordinate system*, E_8 is given by taking the roots in the even coordinate system and changing the sign of any one coordinate. The roots with integer entries are the same while those with half-integer entries have an odd number of minus signs rather than an even number.

Dynkin diagram

The Dynkin diagram for E_8 is given by .

This diagram gives a concise visual summary of the root structure. Each node of this diagram represents a simple root. A line joining two simple roots indicates that they are at an angle of 120° to each other. Two simple roots which are not joined by a line are orthogonal.

Cartan matrix

The Cartan matrix of a rank r root system is an $r \times r$ matrix whose entries are derived from the simple roots. Specifically, the entries of the Cartan matrix are given by

$$A_{ij} = 2\frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}$$

where $(\ ,\ )$ is the Euclidean inner product and α_i are the simple roots. The entries are independent of the choice of simple roots (up to ordering).

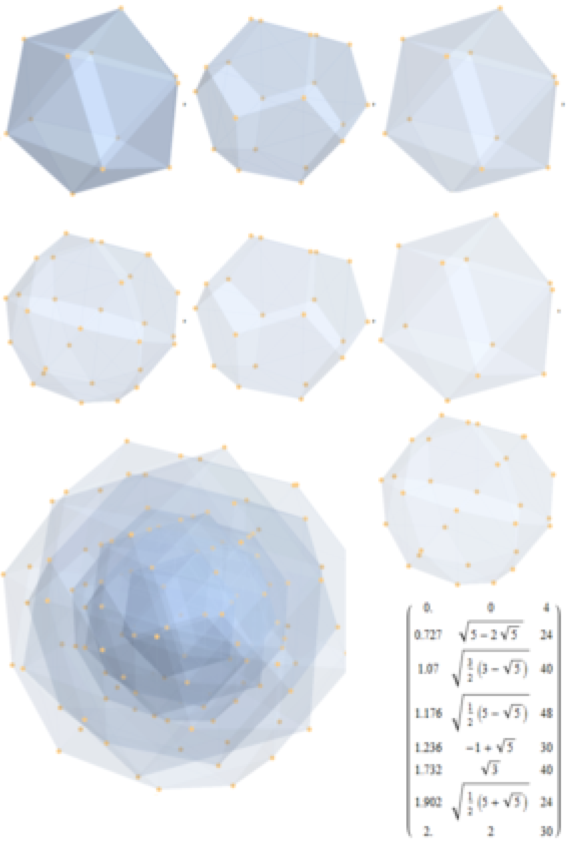
The Cartan matrix for E_8 is given by

$$\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}.$$

The determinant of this matrix is equal to 1.

Simple roots

A set of simple roots for a root system Φ is a set of roots that form a basis for the Euclidean space spanned by Φ with the special property that each root has components with respect to this basis that are either all nonnegative or all nonpositive.



$$\begin{pmatrix} 0. & 0 & 4 \\ 0.727 & \sqrt{5-2\sqrt{5}} & 24 \\ 1.07 & \sqrt{\frac{1}{2}(3-\sqrt{5})} & 40 \\ 1.176 & \sqrt{\frac{1}{2}(5-\sqrt{5})} & 48 \\ 1.236 & -1+\sqrt{5} & 30 \\ 1.732 & \sqrt{3} & 40 \\ 1.902 & \sqrt{\frac{1}{2}(5+\sqrt{5})} & 24 \\ 2. & 2 & 30 \end{pmatrix}$$

Shown in 3D projection using the basis vectors [u,v,w] giving H3 symmetry:

u = $(1, \varphi, 0, -1, \varphi, 0,0,0)$

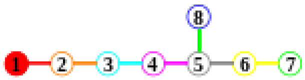
v = $(\varphi, 0, 1, \varphi, 0, -1,0,0)$

w = $(0, 1, \varphi, 0, -1, \varphi,0,0)$

The projected 4_{21} polytope vertices are sorted and tallied by their 3D norm generating the increasingly transparent hulls of each set of tallied norms. These show:

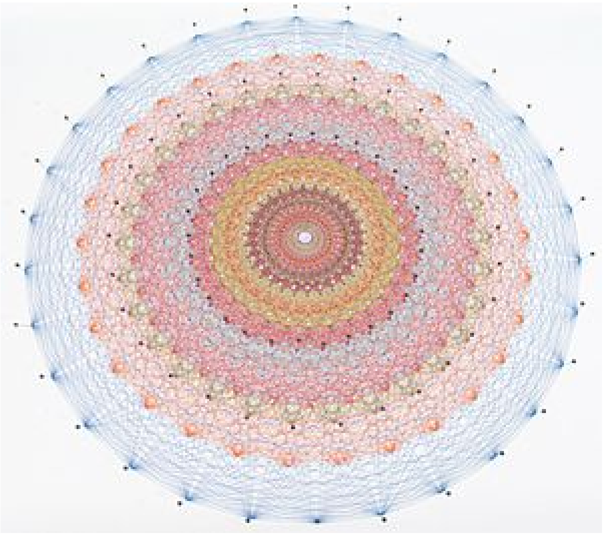
1. 4 points at the origin
2. 2 icosahedrons
3. 2 dodecahedrons
4. 4 icosahedrons
5. 1 icosadodecahedron
6. 2 dodecahedrons
7. 2 icosahedrons
8. 1 icosadodecahedron

for 240 vertices. This is, of course, 2 concentric sets of hulls from the H4 symmetry of the 600-cell scaled by the golden ratio.[2]

Given the E_8 Cartan matrix (above) and a Dynkin diagram node ordering of: 

One choice of simple roots is given by the rows of the following matrix:

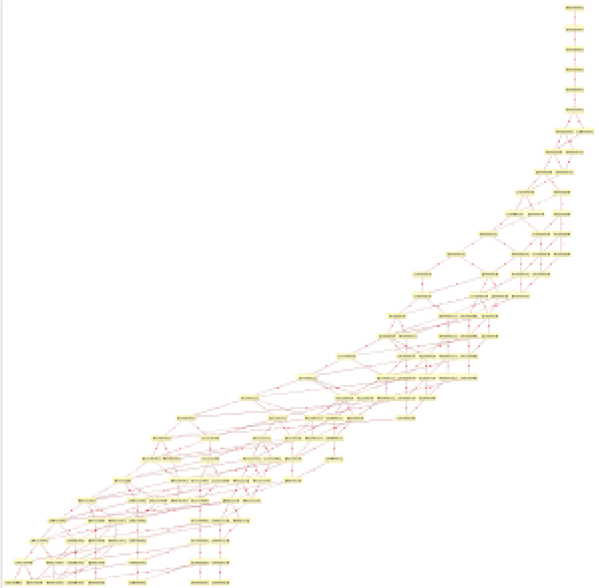
$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}.$$



E8 2d projection with thread made by hand

Weyl group

The Weyl group of E_8 is of order 696729600, and can be described as $O_8^+(2)$: it is of the form 2.*G*.2 (that is, a stem extension by the cyclic group of order 2 of an extension of the cyclic group of order 2 by a group *G*) where *G* is the unique simple group of order 174182400 (which can be described as $PS\Omega_8^+(2)$).[3]



Hasse diagram of E8 root poset with edge labels identifying added simple root position.

E_8 root lattice

The integral span of the E_8 root system forms a lattice in $\mathbf{R}^8$ naturally called the **E_8 root lattice**. This lattice is rather remarkable in that it is the only (nontrivial) even, unimodular lattice with rank less than 16.

Simple subalgebras of E_8

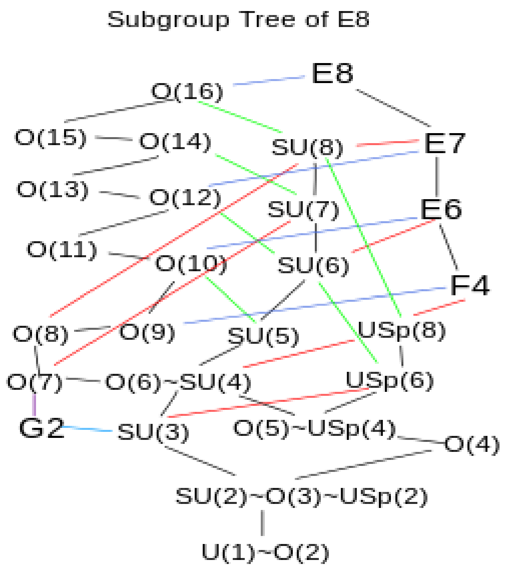
The Lie algebra E8 contains as subalgebras all the exceptional Lie algebras as well as many other important Lie algebras in mathematics and physics. The height of the Lie algebra on the diagram approximately corresponds to the rank of the algebra. A line from an algebra down to a lower algebra indicates that the lower algebra is a subalgebra of the higher algebra.

Chevalley groups of type E_8

Chevalley (1955) showed that the points of the (split) algebraic group E_8 (see above) over a finite field with q elements form a finite Chevalley group, generally written $E_8(q)$, which is simple for any q,[4][5] and constitutes one of the infinite families addressed by the classification of finite simple groups. Its number of elements is given by the formula (sequence A008868 in the OEIS):

$$q^{120}\left(q^{30}-1\right)\left(q^{24}-1\right)\left(q^{20}-1\right)\left(q^{18}-1\right)\left(q^{14}-1\right)\left(q^{12}-1\right)\left(q^{8}-1\right)\left(q^{2}-1\right)$$

The first term in this sequence, the order of $E_8(2)$, namely



An incomplete simple subgroup tree of E_8

337 804 753 143 634 806 261 388 190 614 085 595 079 991 692 242 467 651 576 160 959 909 068 800 000 ≈ 3.38×10^{74}, is already larger than the size of the Monster group. This group $E_8(2)$ is the last one described (but without its character table) in the ATLAS of Finite Groups.[6]

The Schur multiplier of $E_8(q)$ is trivial, and its outer automorphism group is that of field automorphisms (i.e., cyclic of order f if $q=p^f$ where p is prime).

Lusztig (1979) described the unipotent representations of finite groups of type E_8.

Subgroups

The smaller exceptional groups E_7 and E_6 sit inside E_8. In the compact group, both $E_7\times SU(2)/(-1,-1)$ and $E_6\times SU(3)/(\mathbf{Z}/3\mathbf{Z})$ are maximal subgroups of E_8.

The 248-dimensional adjoint representation of E_8 may be considered in terms of its restricted representation to the first of these subgroups. It transforms under $E_7\times SU(2)$ as a sum of tensor product representations, which may be labelled as a pair of dimensions as (3,1) + (1,133) + (2,56) (since there is a quotient in the product, these notations may strictly be taken as indicating the infinitesimal (Lie algebra) representations). Since the adjoint representation can be described by the roots together with the generators in the Cartan subalgebra, we may see that decomposition by looking at these. In this description,

- (3,1) consists of the roots (0,0,0,0,0,0,1,−1), (0,0,0,0,0,0,−1,1) and the Cartan generator corresponding to the last dimension;
- (1,133) consists of all roots with (1,1), (−1,−1), (0,0), $(-\tfrac{1}{2},-\tfrac{1}{2})$ or $(\tfrac{1}{2},\tfrac{1}{2})$ in the last two dimensions, together with the Cartan generators corresponding to the first seven dimensions;
- (2,56) consists of all roots with permutations of (1,0), (−1,0) or $(\tfrac{1}{2},-\tfrac{1}{2})$ in the last two dimensions.

The 248-dimensional adjoint representation of E_8, when similarly restricted, transforms under $E_6\times SU(3)$ as: (8,1) + (1,78) + (3,27) + $(\overline{3},\overline{27})$. We may again see the decomposition by looking at the roots together with the generators in the Cartan subalgebra. In this description,

- (8,1) consists of the roots with permutations of (1,−1,0) in the last three dimensions, together with the Cartan generator corresponding to the last two dimensions;
- (1,78) consists of all roots with (0,0,0), ($-\frac{1}{2}$,$-\frac{1}{2}$,$-\frac{1}{2}$) or ($\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$) in the last three dimensions, together with the Cartan generators corresponding to the first six dimensions;
- (3,27) consists of all roots with permutations of (1,0,0), (1,1,0) or ($-\frac{1}{2}$,$\frac{1}{2}$,$\frac{1}{2}$) in the last three dimensions.
- ($\overline{3}$,$\overline{27}$) consists of all roots with permutations of (−1,0,0), (−1,−1,0) or ($\frac{1}{2}$,$-\frac{1}{2}$,$-\frac{1}{2}$) in the last three dimensions.

The finite quasisimple groups that can embed in (the compact form of) E_8 were found by Griess & Ryba (1999).

The Dempwolff group is a subgroup of (the compact form of) E_8. It is contained in the Thompson sporadic group, which acts on the underlying vector space of the Lie group E_8 but does not preserve the Lie bracket. The Thompson group fixes a lattice and does preserve the Lie bracket of this lattice mod 3, giving an embedding of the Thompson group into $E_8(\mathbf{F}_3)$.

Applications

The E_8 Lie group has applications in theoretical physics and especially in string theory and supergravity. $E_8 \times E_8$ is the gauge group of one of the two types of heterotic string and is one of two anomaly-free gauge groups that can be coupled to the $N = 1$ supergravity in ten dimensions. E_8 is the U-duality group of supergravity on an eight-torus (in its split form).

One way to incorporate the standard model of particle physics into heterotic string theory is the symmetry breaking of E_8 to its maximal subalgebra $SU(3) \times E_6$.

In 1982, Michael Freedman used the E_8 lattice to construct an example of a topological 4-manifold, the E_8 manifold, which has no smooth structure.

Antony Garrett Lisi's incomplete "An Exceptionally Simple Theory of Everything" attempts to describe all known fundamental interactions in physics as part of the E_8 Lie algebra.[7][8]

R. Coldea, D. A. Tennant, and E. M. Wheeler et al. (2010) reported an experiment where the electron spins of a cobalt-niobium crystal exhibited, under certain conditions, two of the eight peaks related to E_8 that were predicted by Zamolodchikov (1989).[9][10]

History

Wilhelm Killing (1888a, 1888b, 1889, 1890) discovered the complex Lie algebra E_8 during his classification of simple compact Lie algebras, though he did not prove its existence, which was first shown by Élie Cartan. Cartan determined that a complex simple Lie algebra of type E_8 admits three real forms. Each of them gives rise to a simple Lie group of dimension 248, exactly one of which (as for any complex simple Lie algebra) is compact. Chevalley (1955) introduced algebraic groups and Lie algebras of type E_8 over other fields: for example, in the case of finite fields they lead to an infinite family of finite simple groups of Lie type.

See also

- E_n

Notes

1. Платонов, Владимир П.; Рапинчук, Андрей С. (1991), *Алгебраические группы и теория чисел*, Наука, ISBN 5-02-014191-7 (English translation: Platonov, Vladimir P.; Rapinchuk, Andrei S. (1994), *Algebraic groups and number theory*, Academic Press, ISBN 0-12-558180-7), §2.2.4
2. The 600-Cell (https://johncarlosbaez.wordpress.com/2017/12/16/the-600-cell)
3. Conway, John Horton; Curtis, Robert Turner; Norton, Simon Phillips; Parker, Richard A; Wilson, Robert Arnott (1985), *Atlas of Finite Groups: Maximal Subgroups and Ordinary Characters for Simple Groups*, Oxford University Press, p. 85, ISBN 0-19-853199-0
4. Carter, Roger W. (1989), *Simple Groups of Lie Type*, Wiley Classics Library, John Wiley & Sons, ISBN 0-471-50683-4
5. Wilson, Robert A. (2009), *The Finite Simple Groups*, Graduate Texts in Mathematics, **251**, Springer-Verlag, ISBN 1-84800-987-9
6. Conway &al, *op. cit.*, p. 235.
7. A. G. Lisi; J. O. Weatherall (2010). "A Geometric Theory of Everything" (http://www.scientificamerican.com/article.cfm?id=a-geometric-theory-of-everything). *Scientific American*. **303** (6): 54–61. Bibcode:2010SciAm.303f..54L (https://ui.adsabs.harvard.edu/abs/2010SciAm.303f..54L). doi:10.1038/scientificamerican1210-54 (https://doi.org/10.1038%2Fscientificamerican1210-54). PMID 21141358 (https://pubmed.ncbi.nlm.nih.gov/21141358).
8. Greg Boustead (2008-11-17). "Garrett Lisi's Exceptional Approach to Everything" (http://seedmagazine.com/news/2008/11/garrett_lisis_exceptional_appr.php). *SEED Magazine*.
9. Most beautiful math structure appears in lab for first time (https://www.newscientist.com/article/dn18356-most-beautiful-math-structure-appears-in-lab-for-first-time.html), *New Scientist*, January 2010 (retrieved January 8, 2010).
10. Did a 1-dimensional magnet detect a 248-dimensional Lie algebra? (http://www.ams.org/notices/201108/rtx110801055p.pdf), *Notices of the American Mathematical Society*, September 2011.

References

- Adams, J. Frank (1996), *Lectures on exceptional Lie groups* (https://books.google.com/books?isbn=0226005275), Chicago Lectures in Mathematics, University of Chicago Press, ISBN 978-0-226-00526-3, MR 1428422 (https://www.ams.org/mathscinet-getitem?mr=1428422)
- Baez, John C. (2002), "The octonions" (http://www.ams.org/bull/2002-39-02/S0273-0979-01-00934-X/home.html), *Bulletin of the American Mathematical Society (N.S.)*, **39** (2): 145–205, arXiv:math/0105155 (https://arxiv.org/abs/math/0105155), doi:10.1090/S0273-0979-01-00934-X (https://doi.org/10.1090%2FS0273-0979-01-00934-X), MR 1886087 (https://www.ams.org/mathscinet-getitem?mr=1886087)
- Chevalley, Claude (1955), "Sur certains groupes simples" (http://projecteuclid.org/euclid.tmj/1178245104), *The Tohoku Mathematical Journal*, Second Series, **7**: 14–66, doi:10.2748/tmj/1178245104 (https://doi.org/10.2748%2Ftmj%2F1178245104), ISSN 0040-8735 (https://www.worldcat.org/issn/0040-8735), MR 0073602 (https://www.ams.org/mathscinet-getitem?mr=0073602)
- Coldea, R.; Tennant, D. A.; Wheeler, E. M.; Wawrzynska, E.; Prabhakaran, D.; Telling, M.; Habicht, K.; Smeibidl, P.; Kiefer, K. (2010), "Quantum Criticality in an Ising Chain: Experimental Evidence for Emergent E_8 Symmetry", *Science*, **327** (5962): 177–180, arXiv:1103.3694 (https://arxiv.org/abs/1103.3694), Bibcode:2010Sci...327..177C (https://ui.adsabs.harvard.edu/abs/2010Sci...327..177C), doi:10.1126/science.1180085 (https://doi.org/10.1126%2Fscience.1180085)
- Garibaldi, Skip (2016), "E_8, the most exceptional group", *Bulletin of the American Mathematical Society*, **53**: 643–671, arXiv:1605.01721 (https://arxiv.org/abs/1605.01721), doi:10.1090/bull/1540 (https://doi.org/10.1090%2Fbull%2F1540)

- Griess, Robert L.; Ryba, A. J. E. (1999), "Finite simple groups which projectively embed in an exceptional Lie group are classified!" (http://www.ams.org/bull/1999-36-01/S0273-0979-99-00771-5/home.html), *American Mathematical Society. Bulletin. New Series*, **36** (1): 75–93, doi:10.1090/S0273-0979-99-00771-5 (https://doi.org/10.1090%2FS0273-0979-99-00771-5), MR 1653177 (https://www.ams.org/mathscinet-getitem?mr=1653177)
- Killing, Wilhelm (1888a), "Die Zusammensetzung der stetigen endlichen Transformationsgruppen" (http://gdz.sub.uni-goettingen.de/index.php?id=11&PPN=GDZPPN002250810&L=1), *Mathematische Annalen*, **31** (2): 252–290, doi:10.1007/BF01211904 (https://doi.org/10.1007%2FBF01211904)
- Killing, Wilhelm (1888b), "Die Zusammensetzung der stetigen endlichen Transformationsgruppen" (http://gdz.sub.uni-goettingen.de/index.php?id=11&PPN=GDZPPN002251337&L=1), *Mathematische Annalen*, **33** (1): 1–48, doi:10.1007/BF01444109 (https://doi.org/10.1007%2FBF01444109)
- Killing, Wilhelm (1889), "Die Zusammensetzung der stetigen endlichen Transformationsgruppen" (https://web.archive.org/web/20150221152955/http://gdz.sub.uni-goettingen.de/index.php?id=11&PPN=PPN235181684_0034&DMDID=DMDLOG_0009&L=1), *Mathematische Annalen*, **34** (1): 57–122, doi:10.1007/BF01446792 (https://doi.org/10.1007%2FBF01446792), archived from the original (http://gdz.sub.uni-goettingen.de/index.php?id=11&PPN=PPN235181684_0034&DMDID=DMDLOG_0009&L=1) on 2015-02-21, retrieved 2013-09-12
- Killing, Wilhelm (1890), "Die Zusammensetzung der stetigen endlichen Transformationsgruppen" (http://gdz.sub.uni-goettingen.de/index.php?id=11&PPN=GDZPPN002252392&L=1), *Mathematische Annalen*, **36** (2): 161–189, doi:10.1007/BF01207837 (https://doi.org/10.1007%2FBF01207837)
- Landsberg, Joseph M.; Manivel, Laurent (2001), "The projective geometry of Freudenthal's magic square", *Journal of Algebra*, **239** (2): 477–512, arXiv:math/9908039 (https://arxiv.org/abs/math/9908039), doi:10.1006/jabr.2000.8697 (https://doi.org/10.1006%2Fjabr.2000.8697), MR 1832903 (https://www.ams.org/mathscinet-getitem?mr=1832903)
- Lusztig, George (1979), "Unipotent representations of a finite Chevalley group of type E8", *The Quarterly Journal of Mathematics. Oxford. Second Series*, **30** (3): 315–338, doi:10.1093/qmath/30.3.301 (https://doi.org/10.1093%2Fqmath%2F30.3.301), ISSN 0033-5606 (https://www.worldcat.org/issn/0033-5606), MR 0545068 (https://www.ams.org/mathscinet-getitem?mr=0545068)
- Lusztig, George; Vogan, David (1983), "Singularities of closures of K-orbits on flag manifolds", *Inventiones Mathematicae*, Springer-Verlag, **71** (2): 365–379, Bibcode:1983InMat..71..365L (https://ui.adsabs.harvard.edu/abs/1983InMat..71..365L), doi:10.1007/BF01389103 (https://doi.org/10.1007%2FBF01389103)
- Zamolodchikov, A. B. (1989), "Integrals of motion and S-matrix of the (scaled) $T=T_c$ Ising model with magnetic field", *International Journal of Modern Physics A*, **4** (16): 4235–4248, Bibcode:1989IJMPA...4.4235Z (https://ui.adsabs.harvard.edu/abs/1989IJMPA...4.4235Z), doi:10.1142/S0217751X8900176X (https://doi.org/10.1142%2FS0217751X8900176X), MR 1017357 (https://www.ams.org/mathscinet-getitem?mr=1017357)

External links

Lusztig–Vogan polynomial calculation

- Atlas of Lie groups (http://www.liegroups.org/)
- Kazhdan–Lusztig–Vogan Polynomials for E_8 (http://www.liegroups.org/kle8.html)
- Narrative of the Project to compute Kazhdan–Lusztig Polynomials for E_8 (http://atlas.math.umd.edu/kle8.narrative.html)

- American Institute of Mathematics (March 2007), *Mathematicians Map E_8* (http://aimath.org/E8/)
- The *n*-Category Café (http://golem.ph.utexas.edu/category/2007/03/news_about_e8.html), a University of Texas blog posting by John Baez on E_8.

Other links

- Graphic representation of E_8 root system (http://www-math.mit.edu/~dav/e8plane.html).
- The list of dimensions of irreducible representations of the complex form of E_8 is sequence A121732 in the OEIS.

Retrieved from "https://en.wikipedia.org/w/index.php?title=E8_(mathematics)&oldid=1013130586"

This page was last edited on 20 March 2021, at 05:15 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.

Sporadic groups and string theory.

First European Congress of Mathematics, Vol. I (Paris, 1992), 411–421, Progr. Math., 119, Birkhäuser, Basel, 1994.

Richard E. Borcherds, DPMMS, 16 Mill Lane, Cambridge, CB2 1SB, England.

This is an expanded version of my talk at the ECM.

1. Some classical infinite product identities.

I will start by giving some well known product identities. The first is

$$\sum_{n\in\mathbf{Z}}(-1)^n q^{3(n+1/6)^2/2} = q^{1/24}\prod_{n>0}(1-q^n).$$

This identity was found by Euler while investigating the partition function; the right hand side is essentially the inverse of the generating function $\prod(1-q^n)^{-1}=\sum p(n)q^n$ of the partition function, and the left hand side is a theta function. A similar identity due to Gauss is

$$\sum_{n\in\mathbf{Z}} q^{n^2} = (1+q)^2(1-q^2)(1+q^3)^2(1-q^4)\dots$$

Both of these are special cases of Jacobi's triple product identity

$$\sum_{n\in\mathbf{Z}}(-1)^n q^{n^2}z^n = \prod_{n>0}(1-q^{2n})(1-q^{2n-1}z)(1-q^{2n-1}z^{-1})$$

if we choose z to be some fixed power of q. Finally Weyl's denominator formula for finite dimensional Lie algebras is

$$e^\rho\sum_{w\in W}\det(w)e^{-w(\rho)} = \prod_{\alpha>0}(1-e^\alpha)$$

where W is the Weyl group, ρ is the Weyl vector, and the product is over all positive roots α. This is a special case of Weyl's character formula which says that the character of a finite dimensional representation is equal to a sum similar to the left hand side divided by the product on the right hand side; for the 1-dimensional representation the character is 1 so the sum is equal to the product.

The Weyl character formula looks rather like the other product formulas. This similarity was explained by Macdonald and Kac as follows. Macdonald observed that the Weyl denominator formula was just a statement about finite root systems, and generalized this to affine root systems, producing a collection of identities called the Macdonald identities. The Macdonald identity for the simplest affine root system is just the Jacobi triple product identity. Kac observed that the Macdonald identities were just the denominator formulas for the Kac-Moody algebras that he and Moody had discovered a few years before, and went on to prove a character formula for representations of these algebras generalizing the Weyl character formula. The Weyl-Kac denominator formula [12] for Kac-Moody algebras is

$$e^\rho\sum_{w\in W}\det(w)e^{-w(\rho)} = \prod_{\alpha>0}(1-e^\alpha)^{\mathrm{mult}(\alpha)}$$

where $\mathrm{mult}(\alpha)$ is the multiplicity of the root α, which is just the dimension of the vector space corresponding to the root α. For finite dimensional Lie algebras the multiplicity of a root is always 1 which is why this expression does not appear in the Weyl denominator formula. For affine Lie algebras such as $SL_n(\mathbf{R}[z,z^{-1}])$ this formula is just a Macdonald identity; for example the Jacobi triple product identity is just the Weyl-Kac denominator formula for $SL_2(\mathbf{R}[z,z^{-1}])$.

I will now briefly describe a proof of the Weyl-Kac character formula using Lie algebra cohomology due to Bott, Kostant, Garland and Lepowsky [10, 14]. Perhaps the best way to motivate Lie algebra cohomology is by analogy with the de Rham cohomology of a compact Riemannian manifold M. The de Rham cohomology of M is defined to be the cohomology groups of the sequence

$$\Lambda^0 \longrightarrow \Lambda^1 \longrightarrow \Lambda^2 \longrightarrow \dots$$

where Λ^n is the vector space of smooth n-forms over M and the arrows are given by the exterior derivative d. The Riemannian manifold on M defines inner products on all the spaces Λ^n which we can use to define the adjoint d^* of d. The Laplace operator is defined as $\Omega = dd^* + d^*d$, and the main theorem of Hodge theory says that the n'th cohomology group H^n is isomorphic to the kernel of Ω on Λ^n. If we have a complex like the one above, the Euler-Poincaré principle states that the alternating sum $\Lambda^* = \Lambda^0 - \Lambda^1 + \Lambda^2 - \ldots$ is equal to the alternating sum $H^* = H^0 - H^1 + H^2 - \ldots$. Strictly speaking, this does not make sense unless the groups Λ^n are all finite dimensional (which they are not) and almost all 0. I will deal with this problem by ignoring it.

Lie algebra cohomology is very similar except that we replace the space of n forms on a manifold by the n'th exterior power of the dual of some Lie algebra, and replace the Riemannian metric on M by an invariant (or contravariant) bilinear form on the Lie algebra. Then, at least for the Lie algebras I will be talking about, the analogues of Hodge theory and the Euler-Poincaré principle are both true.

The Weyl-Kac character formula now turns out to be the Euler-Poincaré principle applied to the cohomology of a suitable Lie algebra. Any Kac-Moody algebra can be written as a (vector space) sum of subalgebras $E \oplus H \oplus F$ where H is the Cartan subalgebra, and E and F are the subalgebras associated to the positive and negative roots. For example, for $SL_n(\mathbf{R})$ the algebras E, F, and H are the upper triangular, lower triangular, and diagonal matrices. The Weyl-Kac denominator formula for the Kac Moody algebra is the Euler-Poincaré principle applied to the cohomology of the subalgebra E, and the cohomology groups can be worked out using Hodge theory. To see this, we use the facts that $\Lambda^*(A \oplus B) = \Lambda^*(A) \otimes \Lambda^*(B)$ for any vector spaces A and B, and $\Lambda^*(A) = 1 - A$ if A is 1-dimensional, so if a vector space is a sum of 1-dimensional spaces A_α, then $\Lambda^*(A)$ is just the product of the factors $(1 - A_\alpha)$. This is where the product in the Weyl-Kac denominator formula comes from. It is more difficult to see why the cohomology gives a sum over the Weyl group; I will just state that the dimension of the n'th cohomology group is the number of elements of the Weyl group of length n. This makes it plausible that the alternating sum over the Weyl group has something to do with the alternating sum of the cohomology groups.

The Weyl-Kac denominator formula is true for non affine Kac-Moody algebras, but unfortunately there are no known examples, other than the finite dimensional and affine algebras, for which the root multiplicities are known explicitly, so this does not really give any new identities. (It is of course always possible to calculate the root multiplicities by using the Weyl-Kac formula or something equivalent to it, but there is no simpler description of them.) However there is a class of Lie algebras, called generalized Kac-Moody algebras [2], which are similar to Kac-Moody algebras, and we can find some examples of these for which the root multiplicities are known. For a finite dimensional Lie algebra all roots have positive norm. For a Kac-Moody algebra, roots can have zero or negative norm, but all simple roots have positive norm, while generalized Kac-Moody algebras can have simple roots of negative norm. Generalized Kac-Moody algebras can also be characterized as Lie algebras with an almost positive definite contravariant bilinear form. More precisely, a Lie algebra G is a generalized Kac-Moody algebra if it satisfies the following conditions:

1. G is graded as $G = \oplus_{n \in \mathbf{Z}} G_n$ such that G_n is finite dimensional for $n \neq 0$.
2. G has an invariant bilinear form compatible with this grading.
3. G has an involution ω preserving $(,)$ and mapping G_n to G_{-n} which acts as -1 on G_0.
4. If we define the contravariant bilinear form $(,)_0$ on G by $(g,h)_0 = (g, \omega(h))$, then $(,)_0$ is positive definite on G_n for $n \neq 0$.

(If we strengthen condition 4 slightly by saying that $(,)_0$ should be positive definite on the whole of G, then the algebras we get are essentially sums of finite dimensional and affine Kac-Moody algebras.)

Many examples of generalized Kac-Moody algebras can be constructed using string theory.

2. String theory.

I will describe how to quantize a string moving in spacetime. The space of states of a quantized string is sometimes a generalized Kac-Moody algebra. See [15] for a more detailed account of the things in this section.

A classical closed string moving in spacetime can be represented by a map from $\mathbf{S}^1 \times \mathbf{R}^1$ to $\mathbf{R}^{n,1}$, where $\mathbf{S}^1$ is the string, $\mathbf{R}^1$ is time on the string, and $\mathbf{R}^{n,1}$ is $n+1$-dimensional spacetime. The classical equations of motion for the string say that the action must be stationary, where the action is some function defined on the space of all maps from $\mathbf{S}^1 \times \mathbf{R}^1$ to $\mathbf{R}^{n,1}$. The original action is the Nambu-Goto action, which is equal to the area of the image of $\mathbf{S}^1 \times \mathbf{R}^1$. (It does not matter that this area is usually infinite, because we are only interested in the change in the area under small local perturbations, and this is well defined. Also it is easier to use a slightly different action called the Polyakov action, but this produces the same solutions to the classical equations of motion.)

Phase space is defined to be the space of all solutions to the classical equations of motion, possibly quotiented out by the action of some group. (It is often defined to be the cotangent space to "configuration space", because this is more or less the same thing, because an element of this cotangent space is roughly the same as giving the position and momentum of everything, and this often determines a unique solution to the classical equations of motion.) The phase space of some classical system can often be made into a symplectic manifold in a canonical way, so that it has a nondegenerate closed 2-form on it. On any symplectic manifold the vector fields preserving the 2-form are called the Hamiltonian vector fields, and can be thought of as the Lie algebra of the group of automorphisms of the symplectic manifold. This Lie algebra has a canonical central extension equal to the Lie algebra of all smooth functions on the manifold under the Poisson bracket (which takes the functions f and g to $\langle df, dg\rangle$). In other words we have the following exact sequence of Lie algebras (at least if the symplectic manifold is compact and simply connected):

$$0 \longrightarrow \mathbf{R} \longrightarrow \Lambda^0 \xrightarrow{d} \Lambda^1_{\text{closed}} \longrightarrow 0,$$

where Λ^0 is the Lie algebra of smooth functions under Poisson bracket, and $\Lambda^1_{\text{closed}}$ is the Lie algebra of closed 1-forms, which are identified with the Hamiltonian vector fields if we use the 2-form to identify 1-forms with vector fields.

The quantization of the string is roughly a representation of some Lie subalgebra of the Lie algebra of functions on phase space. Which subalgebra and which representation we take are up to the person doing the quantization. (This is a gross simplification of what is usually meant by quantization.)

The result of quantizing a (parameterized chiral) string is sometimes a vertex algebra V [1]. Vertex algebras (with some minor extra structure, such as a grading or a bilinear form or an action of the Virasoro algebra) are also called vertex operator algebras [8], meromorphic conformal field theories, W-algebras, and chiral algebras.

To construct the space of states of an unparameterized string, we would like to take the subspace of the vertex algebra V fixed by the diffeomorphism group of the circle $\mathbf{S}^1$, which is $H^0(Diff(\mathbf{S}^1), V)$. This does not work because $Diff(\mathbf{S}^1)$ does not quite act on V; instead we only get a projective action of it on V, so the space of fixed vectors is 0. We can get around this problem by defining the space of physical states of an unparameterized string to be the semi-infinite cohomology group $H^{\infty+1/2}(Vir, V)$ [9]. Here Vir is the Virasoro algebra, spanned by elements L_n, $n \in Z$, and c with the relations

$$\begin{aligned}[L_m, L_n] &= (m-n)L_{m+n} + \delta_m^{-n}(m^3 - m)c/12\\ [L_m, c] &= 0\end{aligned}$$

and is a central extension of the Lie algebra of polynomial vector fields on the circle.

This cohomology group $H^{\infty+1/2}(Vir, V)$ is the generalized Kac-Moody algebra we were trying to construct. It turns out to be nonzero only if spacetime is 26-dimensional; the mysterious number 26 appears because it is minus the value of c when Vir acts on the space of semi-infinite forms over Vir.

3. The monster Lie algebra.

We can apply the denominator formula $H^*(E) = \Lambda^*(E)$ of section 1 to some of the generalized Kac-Moody algebras constructed using string theory in section 2. For one of these algebras, called the monster Lie algebra, this leads to a proof of Conway and Norton's moonshine conjectures for the monster simple group.

The monster Lie algebra [4] is the simplest example of a Lie algebra of physical states of a chiral string on some orbifold. In this case the orbifold is a quotient of the 26-dimensional torus $\mathbf{R}^{25,1}/II_{25,1}$ by an involution, where $II_{25,1}$ is the unique 26-dimensional even unimodular Lorentzian lattice. This algebra can be described in terms of the graded representation $V = \oplus V_n$ of the monster constructed by Frenkel, Lepowsky and Meurman [8], with the property that $\dim(V_n) = c(n-1)$, the coefficient of q^n in the elliptic modular function $j(q) - 744 = \sum c(n)q^n = q^{-1} + 196884q + \ldots$. It follows from the Goddard-Thorn no-ghost theorem [11] that the monster Lie algebra is a $\mathbf{Z}^2$-graded Lie algebra $\oplus_{m,n\in\mathbf{Z}}M_{m,n}$, whose piece $M_{m,n}$ of degree $(m,n)\in\mathbf{Z}^2$ is isomorphic as a module over the monster to V_{mn} if $(m,n)\neq(0,0)$ and to $\mathbf{R}^2$ if $(m,n)=(0,0)$. For small degrees it looks like

$$\begin{array}{ccccccccc}
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & 0 & 0 & 0 & 0 & V_3 & V_6 & V_9 & \cdots \\
\cdots & 0 & 0 & 0 & 0 & V_2 & V_4 & V_6 & \cdots \\
\cdots & 0 & 0 & V_{-1} & 0 & V_1 & V_2 & V_3 & \cdots \\
\cdots & 0 & 0 & 0 & \mathbf{R}^2 & 0 & 0 & 0 & \cdots. \\
\cdots & V_3 & V_2 & V_1 & 0 & V_{-1} & 0 & 0 & \cdots \\
\cdots & V_6 & V_4 & V_2 & 0 & 0 & 0 & 0 & \cdots \\
\cdots & V_9 & V_6 & V_3 & 0 & 0 & 0 & 0 & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

In the decomposition of this algebra as $E\oplus H\oplus F$, the Cartan subalgebra H is the 2-dimensional algebra of degree 0, E is the sum of everything lying to the right of H (in the diagram above), and F is the sum of everything to the left of H. The real roots correspond to the two 1-dimensional spaces V_{-1}, and the simple roots are the ones in the column to the right of H.

To work out the denominator formula $H^*(E)=\Lambda^*(E)$ for E explicitly, we need to know how the Laplace operator Ω acts on $\Lambda^n(E)$. This is easy to describe: it just multiplies any element of degree $(m,n)\in\mathbf{Z}^2$ by $(m-1)n$, so that $H^n(E)$ is just the subspace of elements of $\Lambda^n(E)$ of degree $(1,n)$ or $(m,0)$. This implies that $H^0(E)=\mathbf{R}$, $H^1(E)=\sum_n V_n pq^n$, $H^2(E)=\sum_{m>0}V_m p^{m+1}$, and $H^n=0$ for $n>2$, where we use powers of p and q to keep track of the $\mathbf{Z}^2$ grading.

If we substitute these values into the formula $H^*(E)=\Lambda^*(E)$ and take dimensions of both sides, we obtain the product formula for the j function

$$p^{-1}\prod_{m>0,n\in\mathbf{Z}}(1-p^m q^n)^{c(mn)}=j(p)-j(q)$$

which is the denominator formula of the monster Lie algebra, in the same way that the Macdonald identities are the denominator formulas of the affine Lie algebras.

The monster group acts on the monster Lie algebra by diagram automorphisms. (It is not the full group of diagram automorphisms; its has the special property that $M_{a,b}=M_{c,d}$ as representations of the monster whenever $ab=cd$, $(a,b)\neq(0,0)\neq(c,d)$.) By taking traces of various elements of the monster on the formula $H^*(E)=\Lambda^*(E)$ we obtain the following product formulas involving the coefficients of the Thompson series $T_g(q)=\sum_{n\in\mathbf{Z}}Tr(g|V_n)q^n$:

$$\begin{aligned}&\sum_{m\in\mathbf{Z}}Tr(g|V_m)p^m-\sum_{n\in\mathbf{Z}}Tr(g|V_n)q^n\\ =&p^{-1}\exp(-\sum_{i>0}\sum_{m>0,n\in\mathbf{Z}}Tr(g^i|V_{mn})p^{mi}q^{ni}/i).\end{aligned}$$

For $g=1$ this is just the product formula for the elliptic modular function. These formulas can be rewritten as recursion formulas for the coefficients $Tr(g|V_n)$, and the same recursion formulas for certain modular functions were conjectured by Norton [16] and proved by Koike [13]. This can be used [4] to prove Conway and Norton's moonshine conjectures [5] that the Thompson series $T_g(q)$ are all modular functions (more precisely, normalized Hauptmoduls of genus 0 groups). Norton's generalizations of these conjectures in [17] are still not proved.

4. Modular forms on $O_{n+2,2}(\mathbf{R})$.

The denominator formulas of the examples of many generalized Kac-Moody algebras are modular forms on $O_{n+2,2}(\mathbf{R})$, which is the group of rotations of $\mathbf{R}^{n+2,2}$, and is also the group of conformal transformations of the conformal completion of $\mathbf{R}^{n+1,1}$.

For $n=-1$, $O_{n+2,2}(\mathbf{R})$ is locally isomorphic to $SL_2(\mathbf{R})$ so modular forms on this group are essentially the same as ordinary modular forms. One example is the theta function $\sum q^{n^2}$, which can be written as an infinite product by Gauss's identity.

For $n=0$, $O_{n+2,2}(\mathbf{R})$ is locally isomorphic to $SL_2(\mathbf{R})\times SL_2(\mathbf{R})$. Modular forms on this group are usually Hilbert modular forms, but the examples we give are not, because the discrete subgroup we use is of the form $\Gamma_1\times\Gamma_2$, where each Γ_1 is a subgroup of $SL_2(\mathbf{R})$. (Informally, we could think of them as "degenerate Hilbert modular forms" for the "degenerate real quadratic number field" $Q(\sqrt{1})$.) An example of one of these modular functions is $j(p)-j(q)$, which can be written as an infinite product using the denominator formula of the monster Lie algebra.

For $n=1$, $O_{n+2,2}(\mathbf{R})$ is locally isomorphic to $Sp_4(\mathbf{R})$, so modular forms on this group are essentially Siegel modular forms of genus (or degree) 2. An example of such a form is the Siegel theta function $\sum_{i,j\in\mathbf{Z}} p^{i^2}q^{j^2}r^{ij}$. We will show how to write this as an infinite product in the next section.

I know of modular forms on $O_{n,2}(\mathbf{R})$ which can be written as interesting infinite products for $n=1$, 2, 3, 4, 6, 8, 10, 14, 18, and 26. It is curious that these seem to be exactly the integers for which there is a lacunary modular form of weight $n/2$ which is a product of eta functions (for example $\eta(q)^{26}$); see Serre [18] and Dyson [6]. (The number $n=18$ does not appear in Serre's paper, but the form $\eta(q)^9\eta(q^2)^9$ of weight 18/2 seems to be lacunary.)

5. A superalgebra of rank 3.

I will give a fairly typical example of one of the generalized Kac-Moody superalgebras of higher rank that are (sometimes only conjecturally) the spaces of physical states of some sort of string on some sort of orbifold. Gauss's product identity for the theta function can be written as

$$\sum_{n\in\mathbf{Z}}(-1)^n q^{n^2}=\prod_{n>0}\frac{1-q^n}{1+q^n}.$$

I write it like this to bring out the analogy with the product formula for Siegel's theta function of degree 2, which is

$$\sum_{m,n\in\mathbf{Z}}(-1)^{m+n}p^{m^2}q^{n^2}r^{mn}=\prod_{a+b+c>0}\left(\frac{1-p^aq^cr^b}{1+p^aq^cr^b}\right)^{f(ac-b^2)}$$

where $f(n)$ is defined by $\sum f(n)q^n=1/(\sum_n(-1)^nq^n)=1+2q+4q^2+8q^3+14q^4+\dots$. (The product does not converge for all values of p, q, and r for which the sum is defined.) This is the denominator formula for a generalized Kac-Moody superalgebra of rank 3. This superalgebra is graded by $\mathbf{Z}^3$, and the subspace of degree (a,b,c) had dimension 3 if $(a,b,c)=(0,0,0)$, and $f(ac-b^2)|f(ac-b^2)$ otherwise. (The symbol $m|n$ for the dimension of a superspace means that it is the sum of an ordinary part of dimension m and a super part of dimension n.) This product formula can be proved using the theory of Jacobi forms [7].

The product is a product over all positive semidefinite binary quadratic forms $ax^2+2bxy+cy^2$, and the identity above is equivalent to the following curious fact about such forms. Let $\mathbf{Z}^n$ be a lattice spanned by n pairwise orthogonal vectors of norm 1, and let $r_{a,b,c}(n)$ be the number of representations of the quadratic form $ax^2+2bxy+cy^2$ in the lattice $\mathbf{Z}^n$ which are not contained in any sublattice of the form $\mathbf{Z}^{n-1}$. Then $\sum_{n>0}(-1)^n r_{a,b,c}(n)/n$ depends only on the discriminant and the highest common factor of the coefficients of the form $ax^2+2bxy+cy^2$. This result is unusual because it is usually only possible to say something about the average number of representations of forms of fixed discriminant.

The Mellin transform of Siegel's theta function of degree 2 is essentially a real analytic Eisenstein series (see Zagier [18] for example). Unfortunately it seems to be very difficult to use the product formula for Siegel's theta function to say anything about real analytic Eisenstein series. For that matter, it seems to be just as difficult to use Gauss's product formula for the theta function to say anything about its Mellin transform, the Riemann zeta function.

6. A Lie algebra of rank 26.

I will finish by describing the largest interesting generalized Kac-Moody algebra that I know of [3]. It is the space of physical states of a string moving on a 26-dimensional torus, and is very closely related to the monster Lie algebra, which is the space of states of a string moving on a $\mathbf{Z}_2$ orbifold of the same torus.

The denominator formula for this Lie algebra is

$$e^\rho \prod_{r\in\Pi^+}(1-e^r)^{p_{24}(1-r^2/2)} = \sum_{w\in W}\det(w)w(e^\rho\prod_{n>0}(1-e^{n\rho})^{24}).$$

Here both sides are elements of a completion of the group ring of $II_{1,1}$. The vector ρ is the Weyl vector of this lattice which has inner product -1 with all simple roots of the Weyl group W, and $p_{24}(n)$ is the number of partitions of n into parts of 24 colours.

The Lie algebra itself is graded by $II_{1,1}$, and the piece of degree $r \in II_{1,1}$ has dimension 26 if $r = 0$, and $p_{24}(1-r^2/2)$ otherwise. The algebra is acted on by a group $\Lambda.Aut(\Lambda)$ where Λ is the Leech lattice and $Aut(\Lambda)$ is the double cover of Conway's largest simple group. We can write down a twisted denominator identity similar to the one above for every conjugacy class of this group, by taking traces in the formula $H^*(E) = \Lambda^*(E)$.

The function

$$f(v) = \sum_{w\in W}\det(w)e^{(2\pi i w(\rho),v)}\prod_{n>0}(1-e^{(2\pi i n w(\rho),v)})^{24}.$$

is an automorphic form of the group $O_{26,2}(\mathbf{R})$ with respect to the discrete subgroup $II_{26,2}$. This follows from the functional equation

$$f(2v/(v,v)) = -((v,v)/2)^{12}f(v).$$

I will give a brief sketch of the proof of this functional equation. If we just consider purely imaginary values of v, then f is a solution of the wave equation, and this implies that $((v,v)/2)^{26/2-1}f(2v/(v,v))$ is also a solution by the transformation of the wave operator under the conformal transformation $v \to 2v/(v,v)$. On the other hand, it is easy to show that $f(v)$ vanishes whenever v is imaginary and has norm 2, because the series for $\log(f(v))$ has positive terms and therefore has a singularity on the edge of its region of convergence; this singularity is at all points of the surface C, so f must vanish there because it it regular and $\log(f)$ is not. The fact that f vanishes on this surface easily implies that $f(v)$ and $-((v,v)/2)^{12}f(2v/(v,v))$ both have the same partial derivatives of order at most 1 on this surface. These two functions both satisfy the wave equation and have the same partial derivatives of order at most 1 on a non-characteristic surface, so by the Cauchy-Kovalevsky theorem they must be equal.

The function $f(v)$ also satisfies the trivial functional equations $f(v) = f(v+\lambda)$ for $\lambda \in II_{25,1}$, and $f(w(v)) = \det(w)f(v)$ for w an automorphism of $II_{25,1}$ of spinor norm 1. Together these transformations generate a group isomorphic to the subgroup of index 2 of $Aut(II_{26,2})$ of elements of spinor norm 1. This means that f is essentially a modular form on the group $O_{26,2}(\mathbf{R})$ with respect to the discrete subgroup $Aut(II_{26,2})$. There are many other examples of Lie algebras or superalgebras (for example, two superalgebras of superstrings on a 10 dimensional torus) whose denominator functions are modular forms on $O_{n+2,2}(\mathbf{R})$ for some n. I do not know whether there are an infinite number of such examples.

References.

[1] R. E. Borcherds, Vertex algebras, Kac-Moody algebras, and the monster. Proc. Natl. Acad. Sci. USA. Vol. 83 (1986) 3068-3071.

[2] R. E. Borcherds, Generalized Kac-Moody algebras. J. Algebra 115 (1988), 501–512.

[3] R. E. Borcherds, The monster Lie algebra, Adv. Math. Vol. 83, No. 1, Sept. 1990.

[4] R. E. Borcherds, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math. 109, 405-444 (1992).

[5] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[6] F. Dyson, Missed opportunities, Bull. Amer. Math. Soc. 78 (1972), 635-652.

[7] M. Eichler, D. Zagier, The theory of Jacobi forms, Birkhäuser, 1985.

[8] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[9] I. B. Frenkel, H. Garland, G. Zuckerman, Semi-infinite cohomology and string theory, Proc. Natl. Acad. Sci. U.S.A. 83 (1986), 8442-8446.

[10] H. Garland and J. Lepowsky, Lie algebra homology and the Macdonald-Kac formulas, Invent. math. 34 (1976), p.37-76.

[11] P. Goddard and C. B. Thorn, Compatibility of the dual Pomeron with unitarity and the absence of ghosts in the dual resonance model, Phys. Lett., B 40, No. 2 (1972), 235-238.

[12] V. G. Kac, "Infinite dimensional Lie algebras", third edition, Cambridge University Press, 1990.

[13] M. Koike, On Replication Formula and Hecke Operators, Nagoya University preprint.

[14] B. Kostant, Lie algebra cohomology and the generalized Borel-Weil theorem. Annals of Math. 74, 329-387 (1961).

[15] D. Lüst, S. Theisen. Lectures on string theory. Springer Verlag lecture note in physics 346, 1989.

[16] S. P. Norton, More on moonshine, Computational group theory, Academic press (1984) 185-193.

[17] S. P. Norton, Generalized Moonshine, Proc. Symp. Pure Math. 47 (1987) p. 208-209.

[18] J-P. Serre, Sur la lacunarité des puissances de η, Glasgow Mathematical Journal 27 (1985) 203-221.

[19] D. Zagier, Eisenstein series and the Riemann zeta function, in Automorphic forms, representation theory and arithmetic, Springer Verlag 1981.

"Introduction to Stringy Moonshine "

June, 2018

Tohru Eguchi, University of Tokyo

Monstrous Moonshine

Several years ago we found some curious phenomenon in string theory [1], i.e. appearance of exotic discrete symmetries in the theory. This is now called as moonshine phenomenon and is now under intensive study. Today I would like to give you a brief introduction to moonshine phenomena which may possibly play an interesting role in string theory in the future.

Before going to the moonshine phenomenon in string theory let me briefly recall the story of monstrous moonshine which is well-known. Modular J function has a q-series expansion

$$J(q) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 \\ +20245856256q^4 + 333202640600q^5 + \cdots .$$

$$q = e^{2\pi i\tau},\ \ Im(\tau) > 0,\ \ J(\tau) = J(\frac{a\tau + b}{c\tau + d}),\ \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, Z)$$

It turns out q-expansion coefficients of J-function are decomposed into a sum of dimensions of irreducible representations of the monster group M as

$$196884 = 1 + 196883,\ \ 21493760 = 1 + 196883 + 21296876,$$
$$864299970 = 2 \times 1 + 2 \times 196883 + 21296876 + 842609326,$$
$$20245856256 = 1 \times 1 + 3 \times 196883 + 2 \times 21296876 \\ +842609326 + 19360062527, \cdots .$$

Dimensions of irreducible representations of monster are in fact given by

$$\{1,\ 196883,\ 21296876,\ 842609326,\\ 18538750076,\ 19360062527\cdots\}$$

Monster group is the largest sporadic discrete group, of order $\approx\ 10^{53}$ and the strange relationship between modular form and the largest discrete group was first noted by McKay.

To be precise we may write as

$$J_1(\tau) = J(q) - 744 = \sum_{n=-1} c(n)q^n, \qquad c(0) = 0$$

$$= \sum_{n=-1} Tr_{V(n)}\, 1 \times q^n, \qquad dim V(n) = c(n)$$

McKay-Thompson series is given by

$$J_g(\tau) = \sum_{n=-1} Tr_{V(n)}\, g \times q^n, \quad g \in M$$

where $Tr_{V(n)}\, g$ denotes the character of a group element g in the representation $V(n)$. This depends on the conjugacy class g of M. If McKay-Thompson series is known for all conjugacy classes, decomposition of $V(n)$ into irreducible representations become uniquely determined. Series J_g are modular forms with respect to subgroups of $SL(2, Z)$ and possess similar properties like the modular J-function such as the genus=0 (Hauptmodul) property.

Phenomenon of monstrous moonshine has been understood mathematically in early 1990's using the technology of vertex operator algebra. However, we still do not have a 'simple' physical explanation of this phenomenon.

Elliptic genus

We now consider string theory compactified on K_3 surface. K_3 surface is a complex 2-dimensional hyperKähler manifold and is ubiquitous in string theory. It possesses $SU(2)$ holonomy and a holomorphic 2-form. Thus the string theory on K_3 has an N=4 superconformal symmetry and contains the level $k = 1$ affine $SU(2)$ symmetry and the central charge $c = 6$.

Now instead of modular J-function we consider the elliptic genus of K_3 surface. Elliptic genus describes the topological invariants of the target manifold and counts the number of BPS states in the theory. Using worldsheet variables it is written as

$$Z_{elliptic}(z;\tau) = Tr_{\mathcal{H}_L\times\mathcal{H}_R}(-1)^{F_L+F_R}e^{4\pi izJ^3_{L,0}}q^{L_0-\frac{c}{24}}\bar{q}^{\bar{L}_0-\frac{c}{24}}$$

Here L_0 denotes the zero mode of the Visasoro operators and F_L and F_R are left and right moving fermion numbers. In ellitpic genus the right moving sector is frozen to the supersymmetric ground states (BPS states) while in the left moving sector all the states in the Hilbert space $\mathcal{H}_L$ contribute.

N=4, level k theory contains a SUSY algebra

$$\{\bar{G}^i_0, \bar{G}^{*\,j}_0\} = 2\delta^{ij}\bar{L}_0 - \frac{k}{2}\delta^{ij},\ (i,j=1,2) \Longrightarrow \bar{L}_0 \geq \frac{k}{4}$$

$$\text{BPS states possess} \qquad L_0 = \frac{k}{4}$$

In general it is difficult to compute elliptic genus, however, we were able to evaluate it by making use of Gepner models. Elliptic genus of K3 surface is given by [2]:

$$Z_{K3}(z;\tau) = 8\left[\left(\frac{\theta_2(z;\tau)}{\theta_2(0;\tau)}\right)^2 + \left(\frac{\theta_3(z;\tau)}{\theta_3(0;\tau)}\right)^2 + \left(\frac{\theta_4(z;\tau)}{\theta_4(0;\tau)}\right)^2\right]$$

Here $\theta_i(\tau, z)$ are Jacobi theta functions.

$$Z_{K3}(z=0) = 24, \qquad Z_{K3}(z=\frac{1}{2}) = 16 + O(q),$$
$$Z_{K3}(z=\frac{1+\tau}{2}) = 2q^{-\frac{1}{2}} + O(q^{\frac{1}{2}})$$

It is known that the elliptic genus of a complex D-dimensional manifold is a Jacobi form of weight=0 and index=D/2. When D=2, space of Jacobi form is one-dimensional and given by the above formula.

Jacobi form (weight k and index m) is defined by

$$\varphi(\tau, z + a\tau + b) = e^{-2\pi i m(a^2\tau + 2az)}\varphi(\tau, z),$$
$$\varphi(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}) = (c\tau+d)^k e^{\frac{2\pi i m c z^2}{c\tau+d}}\varphi(\tau, z)$$

We would like to study the decomposition of the elliptic genus in terms of irreducible representations of N=4 SCA. In N=4 SCA, hightest-weight states $|h, \ell\rangle$ are parametrized by

$$L_0|h,\ell\rangle = h|h,\ell\rangle, \qquad J_0^3|h,\ell\rangle = \ell|h,\ell\rangle$$

and the theory possesses two different type of representations, i.e. BPS and non-BPS representations. In the case of $k = 1$ there are representations (in Ramond sector)

$$\text{BPS rep.} \qquad h = \frac{1}{4}; \qquad \ell = 0, \ \frac{1}{2}$$

$$\text{non-BPS rep.} \qquad h > \frac{1}{4}; \qquad \ell = \frac{1}{2}$$

Character of a representation is given by

$$Tr_{\mathcal{R}}(-1)^F q^{L_0} e^{4\pi i z J_0^3}$$

Its index is given by the value at $z = 0$, $Tr_{\mathcal{R}}(-1)^F q^{L_0}$. BPS representations have a non-vanishing index

$$\text{index (BPS, } \ell = 0) = 1$$

$$\text{index (BPS, } \ell = \frac{1}{2}) = -2$$

Character function of $\ell = 0$ BPS representation has the form [3]

$$ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}(z;\tau) = \frac{\theta_1(z;\tau)^2}{\eta(\tau)^3}\mu(z;\tau)$$

where

$$\mu(z;\tau) = \frac{-ie^{\pi i z}}{\theta_1(z;\tau)}\sum_n (-1)^n \frac{q^{\frac{1}{2}n(n+1)}e^{2\pi i n z}}{1 - q^n e^{2\pi i z}}$$

On the other hand the character of non-BPS representations are given by

$$ch^{\tilde{R}}_{h>\frac{1}{4},\ell=\frac{1}{2}} = q^{h-\frac{3}{8}}\ \frac{\theta_1(z;\tau)^2}{\eta(\tau)^3}$$

These have vanishing indices

$$\text{index (non-BPS rep)} = 0$$

At the unitarity bound non-BPS representation splits into a sum of two BPS representations

$$\lim_{h\to\frac{1}{4}} q^{h-\frac{3}{8}}\ \frac{\theta_1^2}{\eta^3} = ch^{\tilde{R}}_{h=\frac{1}{4},\ell=\frac{1}{2}} + 2ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}$$

Function $\mu(z;\tau)$ is a typical example of the so-called Mock theta functions (Lerch sum or Appell function). Mock theta functions look like theta functions but they have anomalous modular transformation laws and are difficult to handle. Recently there were developments in understanding the nature of Mock theta functions due to Zwegers [4]. He has introduced a method of regularization which is similar to those used in physics and improved the modular property of mock theta functions so that they transform as analytic Jacobi forms.

It is possible to derive the following idenities

$$\begin{aligned} ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}(z;\tau) &= \left(\frac{\theta_2(z;\tau)}{\theta_2(0;\tau)}\right)^2 + \mu_2(\tau)\frac{\theta_1(z;\tau)^2}{\eta(\tau)^3} \\ &= \left(\frac{\theta_3(z;\tau)}{\theta_3(0;\tau)}\right)^2 + \mu_3(\tau)\frac{\theta_1(z;\tau)^2}{\eta(\tau)^3} \\ &= \left(\frac{\theta_4(z;\tau)}{\theta_4(0;\tau)}\right)^2 + \mu_4(\tau)\frac{\theta_1(z;\tau)^2}{\eta(\tau)^3} \end{aligned}$$

where

$$\mu_2(\tau)=\mu(z=\frac{1}{2};\tau),\ \mu_3(\tau)=\mu(z=\frac{1+\tau}{2};\tau),\ \mu_4(\tau)=\mu(z=\frac{\tau}{2};\tau)$$

$$\mu(z;\tau) = \frac{-ie^{\pi iz}}{\theta_1(z;\tau)}\sum_n (-1)^n \frac{q^{\frac{1}{2}n(n+1)}e^{2\pi inz}}{1-q^n e^{2\pi iz}}$$

Then we can rewrite the ellitpic genus as

$$Z_{K3}=24ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}(z;\tau)-8\sum_{i=2}^{4}\mu_i(\tau)\frac{\theta_1(z;\tau)^2}{\eta(\tau)^3}$$

Using q-expansion of functions μ_i we find

$$8\left(\mu_2(\tau)+\mu_3(\tau)+\mu_4(\tau)\right)=-2\sum_{n=0}A(n)q^{n-\frac{1}{8}}$$

$$Z_{K3}=24ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}(z,\tau)+2\sum_{n\geq 0}A(n)ch^{\tilde{R}}_{h=\frac{1}{4}+n,\ell=\frac{1}{2}}(z,\tau)$$

At smaller values of n, Fourier coefficients $A(n)$ may be obtained by direct inspection. We find, $A(0)=-1$

n	1	2	3	4	5	6	7	8	$\cdots$
$A(n)$	45	231	770	2277	5796	13915	30843	65550	$\ldots$

Surprize: the dimensions of irreducible reps. of Mathieu group M_{24} appear

$$\text{dimensions}:\{\ 45\quad 231\quad 770\quad 990\quad 1771\quad 2024\quad 2277\quad 3312\quad 3520\quad 5313\quad 5544\quad 5796\quad 10395\quad \cdots\}$$

$$A(6)=13915=3520+10395,$$
$$A(7)=30843=10395+5796+5544+5313+2024+1771$$

Mathieu moonshine [1]

M_{24} is a subgroup of S_{24} (permutation group of 24 objects) and its order is given by $\approx 10^9$. M_{24} is known for its many interesting arithmetic properties and in particular intimately tied to Golay code of error corrections.

Mathieu group appeared before in the work of Mukai on K_3 surface.

Mukai[5] considered K3 surfaces with finite automorphism group. Then these groups are sugbgroups of M_{23}.

♣ Twisted Elliptic Genus

Dimension of a representation equals the trace of the identity representation: we may identify as

$$A(n) = Tr_{V_n}\ 1$$
$$V_1 = 45 + 45^*,\ V_2 = 231 + 231^*,\ V_3 = 770 + 770^*, \cdots$$

We may consider the trace of other group elements in M_{24}

$$A_g(n) = Tr_{V_n}\ g, \qquad\qquad g \in M_{24}$$

$Tr\ g$ depends only on the conjugacy class of g. There exists 26 conjugacy classes $\{g\}$ in M_{24} and also 26 irreducible representations $\{R\}$. We have the character table given by

$$\chi_R{}^g = Tr_R\ g$$

1A	2A	3A	5A	4B	7A	7B	8A	6A	11A	15A	15B	14A	14B	23A	23B	12B	6B	4C	3B	2B	10A	21A	21B	4A
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	7	5	3	3	2	2	1	1	1	0	0	0	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
252	28	9	2	4	0	0	0	1	-1	-1	-1	0	0	-1	-1	0	0	0	0	12	2	0	0	4
253	13	10	3	1	1	1	-1	-2	0	0	0	-1	-1	0	0	1	1	1	1	-11	-1	1	1	-3
1771	-21	16	1	-5	0	0	-1	0	0	1	1	0	0	0	0	-1	-1	-1	7	11	1	0	0	3
3520	64	10	0	0	-1	-1	0	-2	0	0	0	1	1	1	1	0	0	0	-8	0	0	-1	-1	0
45	-3	0	0	1	e_7^+	e_7^-	-1	0	1	0	0	$-e_7^+$	$-e_7^-$	-1	-1	1	-1	1	3	5	0	e_7^-	e_7^+	-3
$\overline{\mathbf{45}}$	-3	0	0	1	e_7^-	e_7^+	-1	0	1	0	0	$-e_7^-$	$-e_7^+$	-1	-1	1	-1	1	3	5	0	e_7^+	e_7^-	-3
990	-18	0	0	2	e_7^+	e_7^-	0	0	0	0	0	e_7^+	e_7^-	1	1	1	-1	-2	3	-10	0	e_7^-	e_7^+	6
$\overline{\mathbf{990}}$	-18	0	0	2	e_7^-	e_7^+	0	0	0	0	0	e_7^-	e_7^+	1	1	1	-1	-2	3	-10	0	e_7^+	e_7^-	6
1035	-21	0	0	3	$2\,e_7^+$	$2\,e_7^-$	-1	0	1	0	0	0	0	0	0	-1	1	-1	-3	-5	0	$-e_7^-$	$-e_7^+$	3
$\overline{\mathbf{1035}}$	-21	0	0	3	$2\,e_7^-$	$2\,e_7^+$	-1	0	1	0	0	0	0	0	0	-1	1	-1	-3	-5	0	$-e_7^+$	$-e_7^-$	3
$\mathbf{1035'}$	27	0	0	-1	-1	-1	1	0	1	0	0	-1	-1	0	0	0	2	3	6	35	0	-1	-1	3
231	7	-3	1	-1	0	0	-1	1	0	e_{15}^+	e_{15}^-	0	0	1	1	0	0	3	0	-9	1	0	0	-1
$\overline{\mathbf{231}}$	7	-3	1	-1	0	0	-1	1	0	e_{15}^-	e_{15}^+	0	0	1	1	0	0	3	0	-9	1	0	0	-1
770	-14	5	0	-2	0	0	0	1	0	0	0	0	0	e_{23}^+	e_{23}^-	1	1	-2	-7	10	0	0	0	2
$\overline{\mathbf{770}}$	-14	5	0	-2	0	0	0	1	0	0	0	0	0	e_{23}^-	e_{23}^+	1	1	-2	-7	10	0	0	0	2
483	35	6	-2	3	0	0	-1	2	-1	1	1	0	0	0	0	0	0	3	0	3	-2	0	0	3
1265	49	5	0	1	-2	-2	1	1	0	0	0	0	0	0	0	0	0	-3	8	-15	0	1	1	-7
2024	8	-1	-1	0	1	1	0	-1	0	-1	-1	1	1	0	0	0	0	0	8	24	-1	1	1	8
2277	21	0	-3	1	2	2	-1	0	0	0	0	0	0	0	0	0	2	-3	6	-19	1	-1	-1	-3
3312	48	0	-3	0	1	1	0	0	1	0	0	-1	-1	0	0	0	-2	0	-6	16	1	1	1	0
5313	49	-15	3	-3	0	0	-1	1	0	0	0	0	0	0	0	0	0	-3	0	9	-1	0	0	1
5796	-28	-9	1	4	0	0	0	-1	-1	1	1	0	0	0	0	0	0	0	0	36	1	0	0	-4
5544	-56	9	-1	0	0	0	0	1	0	-1	-1	0	0	1	1	0	0	0	0	24	-1	0	0	-8
10395	-21	0	0	-1	0	0	1	0	0	0	0	0	0	-1	-1	0	0	3	0	-45	0	0	0	3

Character table of the Mathieu group M_{24} Here we have used $e_p^{\pm} = \frac{1}{2}\left(\pm\sqrt{-p} - 1\right)$.

(Some conjugacy classes are suppressed due to lack of space).

There are two types of conjugacy classes in M_{24}, type I and type II.
Conjugacy class of type I fixes at least one element out of 24 and thus they arise from the conjugacy classes of M_{23}.
On the other hand conjugacy class of type II does not have a fixed point and is intrinsically M_{24}.
For each conjugacy class we want to construct a twisted genus (analogue of McKay-Thompson series in monstrous moonshine)

$$A_g = \sum_{n=1}^{\infty} Tr_{V_n}\; g \times q^n$$

For instance,

$$A_{2A} = -6q + 14q^2 - 28q^3 + 42q^4 - 56q^5 + 86q^6 + \cdots$$

and has the right modular property ($Z_{2A} \in \Gamma_0(2)$).

Twisted genus is decomposed into massless and massive parts

$$Z_g(\tau,z)=\chi_g\, ch^{\tilde{R}}_{h=\frac{1}{4},\ell=0}+\sum_{n\geq 0}A_g(n)ch^{\tilde{R}}_{\frac{1}{4}+n,\ell=\frac{1}{2}}(z,\tau)$$

Here χ_g is the Euler number assigned to the class g

g	1A	2A	3A	5A	4B	7A	8A	6A	11A	15A	14A	23A	others
χ_g	24	8	6	4	4	3	2	2	2	1	1	1	0

χ_g vanishes for type II classes. We note that χ_g can be written as $\chi_g=\chi_1^g+\chi_{23}^g$ which is equal to the number of fixed points of the permutation rep. of g.

conjugacy class	cycle shape
1A	1^{24}
2A	$1^8\cdot 2^8$
3A	$1^6\cdot 3^6$
5A	$1^4\cdot 5^4$
4B	$1^4\cdot 2^2\cdot 4^4$
7A	$1^3\cdot 7^3$
7B	$1^3\cdot 7^3$
8A	$1^2\cdot 2^1\cdot 4^1\cdot 8^2$
6A	$1^2\cdot 2^2\cdot 3^2\cdot 6^2$
11A	$1^2\cdot 11^2$
15A	$1^1\cdot 3^1\cdot 5^1\cdot 15^1$
15B	$1^1\cdot 3^1\cdot 5^1\cdot 15^1$
14A	$1^1\cdot 2^1\cdot 7^1\cdot 14^1$
14B	$1^1\cdot 2^1\cdot 7^1\cdot 14^1$
23A	$1^1\cdot 23^1$
23B	$1^1\cdot 23^1$
12B	12^2
6B	6^4
4C	4^6
3B	3^8
2B	2^{12}
10A	$2^2\cdot 10^2$
21A	$3^1\cdot 21^1$
21B	$3^1\cdot 21^1$
4A	$2^4\cdot 4^4$
12A	$2^1\cdot 4^1\cdot 6^1\cdot 12^1$

Twisted genera for all conjugacy classes of M_{24} have been obtained by our efforts [6, 7]. They reproduce correct lower-order expansion coefficients and are invariant under the Hecke subgroup $\Gamma_0(N)$

$$\Gamma_0(N) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ ad - bc = 1, \ c \equiv 0, \ mod\, N\}$$

N denotes the order of the element g.

From the study of K3 surface with $Z_p(p = 2, 3, \cdots)$ symmetry, for instance, twisted genera of classes $pA(p = 2, 3, \cdots)$ are known [8, 9]

$$Z_{pA}(z;\tau) = \frac{2}{p+1}\phi_{0,1}(z;\tau) + \frac{2p}{p+1}\phi_2^{(p)}(\tau)\phi_{-2,1}(z;\tau)$$

where

$$\phi_{0,1}(z;\tau) = \frac{1}{2}Z_{K3}(z;\tau), \qquad \phi_{-2,1}(z;\tau) = -\frac{\theta_1(z;\tau)^2}{\eta(\tau)^6}$$

are the basis of Jacobi forms with index=1 and

$$\begin{aligned}\phi_2^{(p)}(\tau) &= \frac{24}{p-1}q\partial_q \log\left(\frac{\eta(p\tau)}{\eta(\tau)}\right), \\ &= \frac{24}{p-1}\sum_{k=1}\sigma_1(k)(q^k - pq^{pk})\end{aligned}$$

is an element of $\Gamma_0(p)$.

In the case of type II twisted genera are modular forms of $\Gamma_0(N)$ (with a multiplier system) . They are given in terms of quotients of eta functions.

$$
\begin{aligned}
Z_{2B}(z;\tau) &= 2\frac{\eta(\tau)^8}{\eta(2\tau)^4}\phi_{-2,1}(z;\tau),\\
Z_{3B}(z;\tau) &= 2\frac{\eta(\tau)^6}{\eta(3\tau)^2}\phi_{-2,1}(z;\tau),\\
Z_{4A}(z;\tau) &= 2\frac{\eta(2\tau)^8}{\eta(4\tau)^4}\phi_{-2,1}(z;\tau),\\
Z_{4C}(z;\tau) &= 2\frac{\eta(\tau)^4\eta(2\tau)^2}{\eta(4\tau)^2}\phi_{-2,1}(z;\tau)\\
&\dots
\end{aligned}
$$

n	1A	2A	3A	5A	4B	7A	8A	6A	11A	15A	14A	23A	12B	6B	4C	3B	2B	10A	21A	4A	12A
1	90	-6	0	0	2	-1	-2	0	2	0	1	-2	2	-2	2	6	10	0	-1	-6	0
2	462	14	-6	2	-2	0	-2	2	0	-1	0	2	0	0	6	0	-18	2	0	-2	-2
3	1540	-28	10	0	-4	0	0	2	0	0	0	-1	2	2	-4	-14	20	0	0	4	-2
4	4554	42	0	-6	2	4	-2	0	0	0	0	0	0	4	-6	12	-38	2	-2	-6	0
5	11592	-56	-18	2	8	0	0	-2	-2	2	0	0	0	0	0	0	72	2	0	-8	-2
6	27830	86	20	0	-2	-2	2	-4	0	0	2	0	0	0	6	-16	-90	0	-2	6	0
7	61686	-138	0	6	-10	2	-2	0	-2	0	2	0	-2	-2	-2	30	118	-2	2	6	0
8	131100	188	-30	0	4	-3	0	2	2	0	-1	0	0	0	-12	0	-180	0	0	-4	2
9	265650	-238	42	-10	10	0	-2	2	0	2	0	0	-2	6	10	-42	258	-2	0	-14	-2
10	521136	336	0	6	-8	0	-4	0	0	0	0	2	-2	2	16	42	-352	-2	0	0	0
11	988770	-478	-60	0	-14	6	2	-4	2	0	-2	0	0	0	-6	0	450	0	0	18	0
12	1830248	616	62	8	8	0	0	-2	2	2	0	0	2	-6	-16	-70	-600	0	0	-8	-2
13	3303630	-786	0	0	22	-6	2	0	0	0	-2	2	0	-4	6	84	830	0	0	-18	0
14	5844762	1050	-90	-18	-6	0	2	6	0	0	0	2	0	0	18	0	-1062	-2	0	10	-2
15	10139734	-1386	118	4	-26	-4	-2	6	0	-2	0	0	2	2	-10	-110	1334	4	2	22	-2
16	17301060	1764	0	0	12	0	0	0	-4	0	0	0	2	6	-28	126	-1740	0	0	-12	0
17	29051484	-2212	-156	14	28	0	-4	-4	0	-1	0	0	0	0	12	0	2268	-2	0	-36	0
18	48106430	2814	170	0	-18	8	-2	-6	-2	0	0	-2	2	-6	38	-166	-2850	0	2	14	2
19	78599556	-3612	0	-24	-36	0	0	0	2	0	0	0	-2	-6	-20	210	3540	0	0	36	0
20	126894174	4510	-228	14	14	-6	-2	4	0	2	2	0	0	0	-42	0	-4482	-2	0	-18	0
21	202537080	-5544	270	0	48	4	4	6	-2	0	0	0	-2	6	16	-282	5640	0	-2	-40	2
22	319927608	6936	0	18	-16	-7	4	0	0	0	-1	0	0	4	48	300	-6968	2	-1	24	0
23	500376870	-8666	-360	0	-58	0	-2	-8	4	0	0	2	0	0	-18	0	8550	0	0	54	0
24	775492564	10612	400	-36	28	0	0	-8	0	0	0	0	0	-8	-60	-392	-10556	4	0	-28	-4
25	1191453912	-12936	0	12	64	12	-4	0	0	0	0	0	2	-10	32	462	13064	4	0	-72	0
26	1815754710	15862	-510	0	-34	0	-6	10	0	0	0	-1	0	0	78	0	-15930	0	0	22	-2
27	2745870180	-19420	600	30	-76	-10	4	8	-2	0	-2	0	0	8	-36	-600	19268	-2	2	84	0
28	4122417420	23532	0	0	36	2	0	0	0	0	-2	0	0	12	-84	660	-23460	0	2	-36	0
29	6146311620	-28348	-762	-50	100	-6	4	-10	-2	-2	2	0	0	0	36	0	28548	-2	0	-92	-2
30	9104078592	34272	828	22	-40	0	4	-12	4	-2	0	0	0	-8	96	-840	-34352	-2	0	48	0
31	13401053820	-41412	0	0	-116	0	-4	0	0	0	0	-2	-2	-10	-44	966	41180	0	0	108	0
32	19609321554	49618	-1062	34	50	18	2	10	-2	-2	2	0	0	0	-126	0	-49518	2	0	-46	2
33	28530824630	-59178	1220	0	126	0	-6	12	0	0	0	2	-4	12	62	-1204	59430	0	0	-138	0
34	41286761478	70758	0	-72	-66	-10	-6	0	6	0	2	0	0	12	150	1332	-70890	0	2	54	0
35	59435554926	-84530	-1518	26	-154	6	2	-14	0	2	2	0	0	0	-66	0	84222	2	0	158	2
36	85137361430	100310	1670	0	70	-12	-2	-10	0	0	0	0	-2	-18	-170	-1666	-100170	0	0	-74	-2

♣ Proof of Mathieu moonshine

Orthogonality relation of characters:

$$\sum_g n_g\, \chi_{R'}^g\, \bar{\chi}_R^g = |G|\delta_{RR'}$$

n_g is the number of elements in the conjugacy class g and $|G|$ denotes the order of the group. Let $c_R(n)$ be he multiplicity of representation R in the decompostion of K3 elliptic genus at level n. We then have

$$\sum_R c_R(n)\chi_R^g = A_g(n)$$

Then using the orthogonality relation we find

$$\sum_g \frac{1}{|G|} n_g \bar{\chi}_R^g A_g(n) = c_R(n)$$

We have checked that the multiplicities $c_R(n)$ are all positive integers upto $n = 1000$ and this gives a very strong evidence for Mathieu moonshine conjecture.

n	1	23	252	253	1771	3520	$\frac{\mathbf{45}}{\mathbf{45}}$	$\frac{\mathbf{990}}{\mathbf{990}}$	$\frac{\mathbf{1035}}{\mathbf{1035}}$	$\mathbf{1035'}$	$\frac{\mathbf{231}}{\mathbf{231}}$	$\frac{\mathbf{770}}{\mathbf{770}}$	483	265	2024	2277	3312
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0
8	0	0	0	0	0	2	0	1	1	0	0	0	0	2	0	2	2
9	0	0	0	0	2	4	0	0	2	2	0	2	2	0	2	2	4
10	0	0	0	2	4	8	0	2	2	2	2	0	2	4	4	6	6
11	0	0	0	0	8	12	0	4	4	6	0	4	0	2	10	8	14
12	0	2	2	4	12	30	0	8	8	4	2	6	4	12	12	18	26
13	0	0	4	2	26	44	2	14	14	18	2	10	6	16	30	28	44
14	0	0	4	6	38	86	0	24	24	22	8	16	14	34	46	58	80
15	0	0	12	8	78	144	2	40	44	46	8	38	18	46	86	88	138
16	0	2	18	22	122	252	2	72	72	68	18	50	36	100	140	170	232
17	0	2	30	26	212	410	8	116	124	130	25	94	54	140	246	262	392
18	0	6	50	58	342	704	6	194	202	192	50	148	100	256	388	454	654
19	0	4	80	72	582	1116	18	318	332	346	68	252	150	394	664	722	1062
20	0	14	128	138	904	1836	20	516	536	520	126	390	254	676	1036	1196	1716
21	2	20	214	200	1476	2902	40	814	860	872	182	652	396	1020	1684	1862	2742
22	2	32	328	346	2302	4616	55	1298	1348	1336	314	988	640	1686	2630	3000	4324
23	2	40	512	496	3638	7166	98	2020	2118	2144	460	1590	972	2546	4162	4624	6768
24	0	80	798	824	5584	11192	132	3140	3278	3236	744	2426	1544	4050	6376	7248	10500
25	8	108	1232	1208	8654	17084	234	4814	5038	5084	1106	3764	2336	6108	9892	11042	16112
26	6	174	1860	1904	13090	26148	322	7348	7670	7626	1742	5677	3602	9444	14968	16940	24566
27	12	252	2836	2802	19914	39436	514	11092	11618	11666	2560	8688	5394	14100	22744	25462	37148
28	16	398	4238	4310	29772	59330	742	16686	17418	17356	3922	12912	8160	21414	34026	38434	55764
29	26	560	6328	6286	44512	88280	1154	24840	25994	26078	5758	19380	12090	31636	50892	57068	83146
30	34	876	9368	9486	65776	131020	1642	36824	38480	38368	8642	28580	18008	47172	75158	84776	123176
31	58	1236	13802	13764	97060	192538	2500	54178	56660	56800	12582	42218	26384	69082	110920	124506	181274
32	76	1866	20166	20356	141714	282074	3564	79320	82884	82730	18576	61574	38738	101530	161978	182554	265284
33	122	2664	29396	29374	206524	410062	5286	115334	120644	120798	26830	89868	56226	147156	236010	265136	385974
34	166	3900	42474	42810	298508	593800	7542	166990	174510	174330	39066	129694	81546	213644	341154	384250	558530
35	248	5536	61184	61234	430134	854284	10988	240304	251292	251544	55956	187094	117138	306736	491602	552494	804038
36	334	8058	87622	88196	615626	1224424	15560	344314	359902	359564	80470	267604	168092	440318	703542	792158	1151786

Recently Gannon has proved by mathematical induction that the multiplicities are all positive integers.

Unfortunately the proof so far did not provide much insight into the nature of Mathieu moonshine. The situation is a bit like the case of Monstrous moonshine. 24 of M_{24} will certainly be the Euler number of K_3 and M_{24} permutes homology classes. There are, however, various indications that string theory on K_3 can not have such a high symmetry as M_{24}. Instead of the total Hilbert space the BRS subsector of the theory may possibly possess an enhanced symmetry. It will be interesting to look into the algebraic structures of BPS states to explain Mathieu moonshine.

More Moonshine Phenomena

Recently there have been intense interests in exploring new types of moonshine phenomena other than Mathieu moonshine. Already several types of new moonshine phenomena have been discovered.

- Umbral moonshine [11, 12]
- free fermions on 24 dim. lattice
- moonshine of Spin(7) manifolds
-

. . .

Umbral moonshine was first discovered by generalizing the Mathieu moonshine for the case of Jacobi forms with higher index such as

$$Z(k=2)=a\left[\left(\frac{\theta_2(z)\theta_3(z)}{\theta_2(0)\theta_3(0)}\right)^2+\left(\frac{\theta_2(z)\theta_4(z)}{\theta_2(0)\theta_4(0)}\right)^2+\left(\frac{\theta_3(z)\theta_4(z)}{\theta_3(0)\theta_4(0)}\right)^2\right]$$
$$+b\left[\left(\frac{\theta_2(z)}{\theta_2(0)}\right)^4+\left(\frac{\theta_3(z)}{\theta_3(0)}\right)^4+\left(\frac{\theta_4(z)}{\theta_4(0)}\right)^4\right]$$

When we take a special value $a=4, b=0$, for instance, one finds a moonshine phenomenon with the symmetry group M_{12} acting on this theory.

There is a mysterious relation between Umbral moonshine and Niemeier lattice (self-dual lattices in 24 dimensions). Niemeier lattice is given by a combination of A-D-E type root lattice with appropriate weight vectors so that the lattice becomes self-dual. If one divides the automorphisim group of Niemeier lattice by the automorphism group of A-D-E lattice one obtains discrete groups

$$G_1 = M_{24},\ G_2 = M_{12},\ G_3 = 2.2^3 L_3(2),\ G_4 = 2.S_6,\ G_6 = 2.A_4$$

which agrees exactly with the symmetry groups of Umbral moonshine. At the moment there is no explanation of this coincidence.

Moonshine symmetries recently discovered in string theory are still very mysterious and we may encounter many more surprises in the near future.

References

[1] T.Eguchi, H.Ooguri and Y.Tachikawa, ”Notes on the K3 surface and the Mathiei group M24”, Experimental Mathematics 20 (2011) 91-98.

[2] T.Eguchi, H.Ooguri, A.Taormina and S.-K. Yang, ”Superconformal Algebras and String Compactification on Manifolds with SU(N) Holonomy”, Nucl.Phys.B315 (1989) 193-221.

[3] T.Eguchi and A.Taormina, ”Character formulas for the N=4 superconformal algebra,” Phys.Lett. B 200 (1989) 315.

[4] S.P.Zwegers, ”Moch Theta Functions,” Ph.D. Thesis, arXiv:0807.4834[math.NT].

[5] S.Mukai, ”Finite group of automorphisms of K3 surfaces and the Mathieu group,” Invent. Math. 94 (1988) 183.

[6] M.Gaberdiel, S.Hohenegger and R.Voltapo, ”Mathieu moonshine in the elliptic genus of K3,” JHEP 1010 (2010) 062, arXiv:1008.3778 [hep-th].

[7] T.Eguchi and K.Hikami, ”Note on twisted elliptic genus of K3 surface,” Phys.Lett.B694 (2011) 446,arXiv:1008.4924 [hep-th].

[8] A.Sen, ”Discrete information from CHL black holes”, arXiv;1002.3857[hep-th],2016.

[9] M.Cheng, ”K3 surface, N=4 Dyons, and the Mathieu group M24”; arXiv:1005.5415[hep.th].

[10] T.Gannon, ”Much Ado about Mathieu,” arXiv:1211.5531[math.RT].

[11] M.Cheng, J.Duncan and J.Harvey, ”Umbral moonshine”, Commun.Num.Theor.Phys.08 (2014) 101; ”Umbral moonshine and the Niemeier lattice”, arXiv:1307.5793[math.RT].

[12] M.Cheng and S.Harrison, Umbral moonshine and K3 surface, Commun. Math. Phys. 339 no.1, 221 (2015) [arXiv:1406.0619][hep-th]]

arXiv:hep-th/0001158v2 18 Feb 2000

Finite Group Modular Data

Antoine Coste
CNRS[*] Laboratory of Theoretical Physics
building 210, University of Orsay
91405 Orsay cedex, France
`Antoine.Coste@th.u-psud.fr`

Terry Gannon
Department of Mathematical Sciences,
University of Alberta
Edmonton, Canada, T6G 2G1
`tgannon@math.ualberta.ca`

Philippe Ruelle[†]
Institut de Physique Théorique
Université catholique de Louvain
B-1348 Louvain-La-Neuve, Belgium
`ruelle@fyma.ucl.ac.be`

Abstract

In a remarkable variety of contexts appears the modular data associated to finite groups. And yet, compared to the well-understood affine algebra modular data, the general properties of this finite group modular data has been poorly explored. In this paper we undergo such a study. We identify some senses in which the finite group data is similar to, and different from, the affine data. We also consider the data arising from a cohomological twist, and write down, explicitly in terms of quantities associated directly with the finite group, the modular S and T matrices for a *general* twist, for what appears to be the first time in print.

---

[*] UMR 8627

[†] Chercheur qualifié FNRS

1 Mathematical and Physical Origin

Throughout this paper, let G denote any finite group (good references to finite group and character theory are provided by [1, 2, 3]).

Physically, the modular data we will describe in the next section arise in several ways. It is a (2+1)-dimensional Chern-Simons theory with finite gauge group G [4, 5]. As Witten showed, 3-dimensional topological field theory corresponds to 2-dimensional conformal field theory (CFT), and the corresponding CFTs here are orbifolds by symmetry group G of a holomorphic CFT [6, 4, 7, 8] (e.g. the E_8 level 1 WZW orbifolded by any finite subgroup of the compact simply-connected Lie group $E_8(\mathbb{R})$). This CFT incarnation is important to us, as it provides some motivation for specific investigations we will perform. Nevertheless, both of these incarnations are probably of direct value only as toy models.

Note that more generally, to each G we will obtain a finite-dimensional representation of the mapping class group $\Gamma_{g,n}$ for the genus g surfaces with n punctures [9]. In this paper though we will consider only the modular group $\mathrm{SL}_2(\mathbb{Z})$, corresponding to $\Gamma_{1,0}$, and in particular the matrices S and T.

It is clear that many different CFTs can realise the same modular data – e.g. all 71 of the $c = 24$ holomorphic theories have $S = T = (1)$. So any given G will correspond to several different CFTs. Likewise, all we can say about the central charge c associated to G is that it will be a positive multiple of 8.

Associated with the RCFT we expect to have some sort of quantum-group which captures the modular data and representation theory of the chiral algebra (so e.g. the fusion coefficients are tensor product coefficients of irreducible modules). This was done for arbitrary finite G in [7], and is the quantum-double of the group algebra $\mathbb{C}[G]$ (see also [10, 11, 12]).

There should also be a vertex operator algebra (VOA) interpretation to this data, assigning a VOA to each finite group. One way to do this is to start with the VOA associated with any even self-dual Euclidean lattice Λ for which $G \leq \mathrm{Aut}\,\Lambda$. Orbifolding it by G should yield our data (with the value $c = \dim \Lambda$). For example, any finite subgroup of $\mathrm{SU}_2(\mathbb{C})$ or $\mathrm{SU}_3(\mathbb{C})$ works with the $\Lambda = E_8$ root lattice. By Cayley's Theorem, such a lattice Λ can always be found for a given group G: e.g. embed G in some $\mathfrak{S}_n$ and take Λ to be the orthogonal direct sum of n copies of E_8. In [13, 10], this VOA interpretation is addressed.

Although these holomorphic orbifolds are perhaps too artificial to be of direct interest, it can be expected that they provide a good hint of the behaviour of more general orbifolds. Indeed this is the case – e.g. they can be seen in the theory of permutation orbifolds [14].

Perhaps the most physical incarnation of this modular data is in (2+1)-dimensional quantum field theories where a continuous gauge group has been spontaneously broken to a finite group [15]. Non-abelian anyons (i.e. particles whose statistics are governed by the braid group rather than the symmetric group) arise as topological excitations. The effective field theory describing the long-distance physics is governed by the quantum-group of [7]. Adding a Chern-Simons term corresponds to the cohomological twist to be discussed shortly.

Actually, this modular data arose originally in mathematics, in an important but technical way in Lusztig's determination of the irreducible characters of the finite groups of Lie type [16, 17]. In describing some of these, the so-called 'unipotent' characters, he was led to consider this modular data for the groups $G \in \{\mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2\, , \mathfrak{S}_3\, , \mathfrak{S}_4\, , \mathfrak{S}_5\, \}$. For example, $\mathfrak{S}_5$ arises in groups of type E_8. Our primary fields Φ parametrise the unipotent characters associated to a given 2-sided cell in the Weyl group. [16] interprets the fusion algebra as the Grothendieck ring for G-equivariant vector bundles. This Lusztig interpretation is significant as it indicates the richness of the purely group theoretic side of the data we explore below.

An intriguing application to quantum computing[‡] has been suggested by Kitaev [18]. One of the main challenges of actually implementing an effective quantum computer is decoherence: interaction with the environment makes quantum superpositions very unstable. The standard approaches to this involve quantum error correction, but Kitaev's proposal is to incorporate into the hardware the non-abelian anyons of [15]. The resulting computer would operate quite robustly amidst localised disturbances.

From many of these approaches, we also expect to have characters $\mathrm{ch}_a(\tau)$ which realise this modular data as in (2.1) and (2.2) below. These appear to be given in [8], at least when G is a subgroup of $E_8(\mathbb{C})$ (see Theorem 4.3 there).

In summary, what we get is a set of modular data (i.e. matrices S and T) for any choice of finite group G. Now, much information about a group can be recovered easily from its character table: e.g. whether it is abelian, simple, solvable, nilpotent, ... (see e.g. Ch.22 of [1]). For instance, G is simple iff for all irreducible $\chi \not\equiv 1$, $\chi(a) = \chi(e)$ only for $a = e$. Thus it can be expected that our modular data, which probably includes the character table, should tell us a lot about the group, i.e. be sensitive to a lot of the group-theoretic properties of G.

It does not appear to be known yet whether the modular data associated to distinct G will always be distinct. A tempting guess would be that this modular data is a function of the group algebra, but that it succeeds in distinguishing groups which have different character tables. An easy result along these lines is: Given any groups G and H with equal matrix S, if G is abelian then $G \cong H$. (The character table of an abelian group also uniquely determines it.) The non-abelian order-8 groups $\mathfrak{D}_4$ and $\mathfrak{Q}_4$ have identical character tables, but their matrices S and T are both different. More generally, we will see below that the matrices S for the groups with at most 50 primary fields are all distinct.

There are two ways to generalise this data. One is by a 'twisting' by some element of a cohomology group [4] (see also [6]). We will look at this twisting in Section 5. This twisting is rather interesting, because completely analogously one can speak of the twisting of modular data associated to Lie groups: the cohomology group there is $\mathbb{Z}$, and 'twisting' by cohomology class $[k]$ gives WZW at level k for that Lie group ! So the untwisted finite group modular data is 'level 0' in this sense. On the other hand, the WZW models at level 0 are trivial. Another difference with WZW is that the cohomology group here is finite, so there are only finitely many different possible 'levels' to this finite group data.

The twisting was incorporated in the quantum-group picture in [7]. How to obtain topological (e.g. oriented knot) invariants from this twisted data was explained in [20]. It turns out that these knot invariants are functions of the "knot group" (i.e. the fundamental group of the complement of the knot). Though non-isotopic knots can have the same knot group, this does not imply that these finite group topological invariants are uninteresting. For instance [21], these invariants can distinguish a knot from its inverse (i.e. the knot with opposite orientation), unlike the more familiar topological invariants coming from affine algebras (or quantum groups $U_q(\mathfrak{g})$).

This twisting is reminiscent, though independent, of *discrete torsion* [19]. The twist changes the modular data, while discrete torsion changes the modular invariant partition function but leaves unchanged the modular data. We use discrete torsion in §4.1.

[‡] We thank Sander Bais for informing us of this.

A second way to generalise this data is suggested by the subfactor interpretation of RCFT. There is a von Neumann subfactor and from this a fusion ring, associated to any group-subgroup pair $G \leq H$ [22]. Our data comes from the diagonal embedding of G in $G \times G$. The fusion ring for generic pairs $G \leq H$ is non-commutative (so e.g. Verlinde's formula will not work), but it is commutative for much more than just $G \leq G \times G$ and so we can expect here a significant generalisation of this finite group data corresponding in some way to physics. To our knowledge, such a physical interpretation has not been developed.

The finite group modular data behaves very differently from the affine modular data. Yet finite group theory is certainly richer than nontwisted affine Lie theory, so its modular data should be explored. It should be interesting both from the group theory and RCFT sides. In particular, it should be interesting to RCFT for the same reasons as the affine data: Affine modular data is regarded as physically interesting primarily because it serves as a toy model, and because of the GKO coset construction; similarly, the finite group modular data is equally a toy model, and is involved in the orbifold construction.

As always, there are unexpected surprises. One of the differences we have found between the finite group data and the affine data is the large number of modular invariants for the former – in contrast the affine modular invariants are surprisingly rare and generally are associated with Dynkin diagram symmetries. There is one family of exceptions to this rule: the orthogonal algebras at level 2 have a rich collection of modular invariants [33]. This was rather mysterious. But what we find is that the only overlap between the affine data and nonabelian finite group data occurs with the orthogonal algebras at level 2 ! This clearly provides a partial resolution of that mystery.

Our main intention with this paper is to make the mathematical physics community more familiar with this modular data, by developing and illustrating some of the basic ideas and constructions. As a large and relatively unexplored class of RCFTs, studying it can correct some of the prejudices we have developed with our rather artificial preoccupation with the affine data. Finally, this material suggests the thought: "There's some kind of analogy relating affine algebras and finite groups" – after all, both are directly associated to topological groups – and we try to work out some of its aspects. See the Table below. Of course there are differences as well as similarities, and identifying these will help us develop a better understanding of generic modular data.

List of Notation

G	a group with finitely many elements
$\mathbb{N}$	the nonnegative integers
ch_a	RCFT characters, see e.g. (2.1), (2.8)
ξ_m	the m-th root of unity $\exp[2\pi\mathrm{i}/m]$
$\mathbb{Z}_n$	the cyclic group (the integers modulo n)
$\mathfrak{S}_n$	the symmetric group of order $n!$
$\mathfrak{A}_n$	the alternating group of order $n!/2$

$\mathfrak{D}_n$	the (binary) dihedral group (3.3) of order $2n$
$\mathfrak{Q}_{2n}$	the (generalised) quaternion group (3.4) of order $4n$
e	the identity element in the group
$\|G\|$	the order (cardinality) of G
K_a	the conjugacy class of a, i.e. $\{gag^{-1} \mid g \in G\}$
R	a set of representatives for each conjugacy class K_a
$\chi, \psi, \ldots$	irreducible characters
$\mathrm{Irr}(G)$	the set of all irreducible characters of G
$G(a,b)$	see below (2.12)
$\deg(\chi)$	dimension of corresponding representation
$C_G(a)$	the centraliser of a in G, i.e. $\{g \in G \mid ga = ag\}$
Φ	set of "primary fields" (a, χ) where $\chi \in \mathrm{Irr}(C_G(a))$
$e(G)$	the exponent of G (smallest n such that $g^n = e$ for all $g \in G$)
$Z(G)$	the centre of G
G'	the commutator subgroup $\langle ghg^{-1}h^{-1} \rangle$
$k(G)$	the class number of G, i.e. the number of conjugacy classes
$H \leq G$	H is a subgroup of G
$H \lhd G$	H is a normal subgroup of G
χ_H^G	the induced character of $\chi \in \mathrm{Irr}(H)$, where $H \leq G$
$M(G)$	the Schur multiplier $H^2(G, U(1))$
β	a $U(1)$-valued 2-cocycle
$r(G, \beta)$	the number of inequivalent projective β-representations of G
β-$\mathrm{Irr}(G)$	the set of all irreducible projective β-characters of G
$\tilde{\chi}, \tilde{\psi}, \ldots$	irreducible projective characters
α	a 3-cocycle in $H^3(G, U(1))$
CT	a 3-cocycle α such that all 2-cocycles β_a are coboundaries (see §5.2)
S^α, T^α	the modular data twisted by a 3-cocycle α
$\mathbb{F}_q$	the finite field with q elements

The affine algebra versus finite group analogy

primary fields in Φ	r-tuples $\lambda \in P_+(X_{r,k})$	pairs (g, χ)
level	a number $k \in \mathbb{N}$	a twist $\alpha \in H^3(G, U(1))$
simple currents	extended Dynkin diagram symmetry (except $\hat{E}_{8,2}$)	$Z(G) \times G/G'$ (untwisted case)
conjugation	unextended Dynkin symmetry	involutive Galois permutations
conformal embedding	$X_{r,k} \subset Y_{s,1}$	perhaps $H \lhd G$
central charge	$c = \frac{k \dim X_r}{k + h^\vee}$	$c \in 8\mathbb{Z}$
Chern-Simons theory	gauge group is Lie gp	gauge group is finite group
CFT	WZW model	holomorphic orbifold
quantum-group	$U_q(X_r)$, $q^N = 1$	quantum double $D^\alpha(G)$
VOA	Frenkel-Zhu	Dong-Mason
rank-level duality	$\widehat{\mathrm{su}}(n)_k \leftrightarrow \widehat{\mathrm{su}}(k)_n, \ldots$	electric/magnetic duality(??)

2 Untwisted modular data

2.1 Modular data for any RCFT

In this paper we explore what we call the modular data of G, from the perspective of RCFT. This subsection is intended as a quick review of the basic RCFT data (see e.g. [23] for a more comprehensive treatment), and we end it with a brief description of the modular data associated with affine algebras.

Let Φ denote the (finite) set of primary fields in the RCFT. One of these primaries, the "vacuum" '0', is privileged. By the *modular data* we mean the modular matrices S and T, whose entries S_{ab} and T_{ab} are parametrised by $a, b \in \Phi$. These matrices are unitary, T is diagonal, S is symmetric, and together they define a representation of $\mathrm{SL}_2(\mathbb{Z})$ realised by the modular action on the RCFT characters $\mathrm{ch}_a(\tau)$:

$$\mathrm{ch}_a(-1/\tau) = \sum_{b\in\Phi} S_{ab}\,\mathrm{ch}_b(\tau), \tag{2.1}$$

$$\mathrm{ch}_a(\tau+1) = \sum_{b\in\Phi} T_{ab}\,\mathrm{ch}_b(\tau), \tag{2.2}$$

so $(ST)^3 = S^2 =: C$ commutes with S and T and is a permutation matrix called *charge conjugation.* Note that $S^* = SC$ and $C0 = 0$. These ch_a are assumed to be linearly independent and in general will depend on other variables than τ, but it is conventional to write only τ.

All RCFTs in this paper are *unitary.* This implies that the entries $S_{a0} = S_{0a}$ will all be strictly positive. The ratio $\frac{S_{a0}}{S_{00}}$ is called the *quantum dimension* of a. The *fusion coefficients* $N_{ab}^c \in \mathbb{N}$ of the theory can be obtained from S using *Verlinde's formula*

$$N_{ab}^c = \sum_{d\in\Phi} \frac{S_{ad}\,S_{bd}\,S^*_{cd}}{S_{0d}}. \tag{2.3}$$

A *simple current* [24] can be defined as any $j \in \Phi$ with quantum-dimension 1: i.e. $S_{j0} = S_{00}$. To any simple current j is associated a "charge" $Q_j : \Phi \to \mathbb{Q}$, an integer R_j, and a permutation J of Φ, such that $J0 = j$,

$$S_{Ja,b} = \exp[2\pi\mathrm{i}\, Q_j(b)]\, S_{ab}, \tag{2.4}$$

$$T_{Ja,Ja}T^*_{aa} = \exp[2\pi\mathrm{i}\,(R_j\,\tfrac{n-1}{2n} - Q_j(a))], \tag{2.5}$$

$$N_{j,b}^c = \delta_{c,Jb}, \tag{2.6}$$

where n in (2.5) is the order of J. Note that $nQ_j \in \mathbb{Z}$. For instance $j = 0$ is a simple current, corresponding to the identity permutation. Simple currents define an abelian group given by composition of the corresponding permutations.

The matrix S also obeys another important symmetry, called *Galois* [25]. In particular, the entries S_{ab} will lie in some cyclotomic field $\mathbb{Q}(\xi_m)$, where ξ_m throughout this paper denotes the mth root of unity $\exp[2\pi\mathrm{i}/m]$. This means that each S_{ab} can be written as a polynomial in ξ_m with rational coefficients. The Galois group $\mathrm{Gal}(\mathbb{Q}(\xi_m)/\mathbb{Q})$ consists of all automorphisms of the field $\mathbb{Q}(\xi_m)$, and is isomorphic to the multiplicative group $\mathbb{Z}_m^\times$ of integers coprime to m. In particular, to any $\ell \in \mathbb{Z}_m^\times$ we get an automorphism $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_m)/\mathbb{Q})$ sending ξ_m to ξ_m^ℓ. Applying it to entries of S, we get

$$\sigma(S_{ab}) = \epsilon_\sigma(a)\, S_{\sigma(a),b} = \epsilon_\sigma(b)\, S_{a,\sigma(b)}, \tag{2.7}$$

where each $\epsilon_\sigma(a)$ is a sign ± 1, and where $a \mapsto \sigma(a)$ is a permutation of Φ. The simplest example of this Galois action is charge conjugation C: σ_{-1} is complex conjugation; the corresponding $\epsilon_{\sigma_{-1}}$ is identically $+1$, while the permutation $a \mapsto \sigma_{-1}a$ is given by C.

The Galois and simple current permutations respect each other: $\sigma_\ell(Ja) = J^\ell(\sigma_\ell(a))$. Also, $\epsilon_\sigma(Ja) = \epsilon_\sigma(a)$ and $Q_j(\sigma_\ell(a)) \equiv \ell\, Q_j(a) \pmod 1$.

The one-loop partition function $\mathcal{Z}(\tau)$ of the RCFT is a modular invariant sesquilinear combination of RCFT characters:

$$\mathcal{Z}(\tau) = \sum_{a,b\in\Phi} M_{ab}\,\mathrm{ch}_a(\tau)\,\mathrm{ch}_b(\tau)^*. \tag{2.8}$$

Modular invariance means that the coefficient matrix M commutes with S and T. In addition, $M_{00} = 1$, and each coefficient is a non-negative integer: $M_{ab} \in \mathbb{N}$. Any such $\mathcal{Z}$ or M is called a *physical invariant* or *modular invariant* (such an M may or may not be realised as a partition function for an RCFT). $M = I$ is always a physical invariant. For a given choice of modular data, there will only be finitely many physical invariants.

A special family of physical invariants are the *automorphism invariants*, where M is a permutation matrix. An example is $M = C$. The product MM' of any automorphism invariant M with any physical invariant M' will also be a physical invariant.

One of the uses of both simple currents and this Galois action is that they can be used to construct physical invariants. The simplest construction (originally due to [26] but since generalised considerably, see [27] and references therein) starts with any simple current j, with order n say. Then

$$M(j)_{ab} = \sum_{i=1}^{n} \delta_{J^i a,b}\,\delta^1(Q_j(a) + \frac{i}{2n}R_j), \tag{2.9}$$

where $\delta^1(x) = 1$ if $x \in \mathbb{Z}$ and $= 0$ otherwise. For instance the choice $j = 0$ yields the identity matrix $M(0) = I$. $M(j)$ will be a physical invariant iff $T_{jj}T_{00}^*$ is an nth root of 1. $M(j)$ will be an automorphism invariant iff $T_{jj}T_{00}^*$ is a *primitive* nth root of 1.

Given any order-two Galois automorphism σ with the properties that $\sigma(0) = 0$ and $T_{\sigma a,\sigma a} = T_{a,a}$, then $M(\sigma)_{ab} := \delta_{b,\sigma a}$ defines an automorphism invariant (this construction was originally due to [28] but was since generalised). Again, C is an example.

The most familiar example of modular data comes from the affine nontwisted algebras. The literature on affine modular data is very extensive (indeed this affine ↔ finite group imbalance is a primary motivation for this paper) and it certainly is not our intention to review it here. See [29, 23] for a deeper treatment.

Choose any finite-dimensional simple Lie algebra X_r, and any *level* $k \in \mathbb{N}$. $X_r^{(1)} = \hat{X}_r$ denotes the infinite-dimensional nontwisted affine Kac-Moody algebra [29]. Its integrable highest-weight modules at level k are parametrised by $(r+1)$-tuples $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_r)$, where each $\lambda_i \in \mathbb{N}$, and $\sum_{i=0}^{r} a_i^\vee \lambda_i = k$. The positive constants $a_i^\vee$ are called colabels – e.g. for $A_r^{(1)}$ or $C_r^{(1)}$ all $a_i^\vee = 1$. The (finite) collection of these highest-weights is denoted P_+^k and is the set of primary fields Φ. The "vacuum" 0 is $(k, 0, 0, \ldots, 0)$. The matrices S and T are given in Theorem 13.8 of [29]: S is related to the corresponding Lie group characters at elements of finite order, while T is related to the eigenvalues of the quadratic Casimir. A useful parameter is the dual Coxeter number $h^\vee := \sum_{i=0}^r a_i^\vee$, and a useful $(r+1)$-tuple is the Weyl vector $\rho := (1,1,1,\ldots,1)$. The affine fusion coefficients are given combinatorially by what is usually called the Kac-Walton formula [29, 30], though it has other co-discoverers.

The simplest example is the choice $X_r = A_1$ (i.e. $\mathrm{sl}_2(\mathbb{C})$) for which $h^\vee = 2$. For any level $k \in \mathbb{N}$, one has $P_+^k = \{(k,0),(k-1,1),\ldots,(0,k)\}$. The modular data is then

$$\begin{aligned} S_{\lambda\mu} = & \sqrt{\tfrac{2}{k+2}}\,\sin(\pi\tfrac{(\lambda_1+1)(\mu_1+1)}{k+2}), && (2.10)\\ T_{\lambda\mu} = & \exp[\pi \mathrm{i}\tfrac{(\lambda_1+1)^2}{2(k+2)} - \tfrac{\pi\mathrm{i}}{4}]\,\delta_{\lambda,\mu}. && (2.11)\end{aligned}$$

The charge conjugation of affine modular data corresponds to an order-one or -two symmetry of the unextended Dynkin diagram. For $A_1^{(1)}$ it is trivial, but for the other $A_r^{(1)}$ it is non-trivial. The simple currents were classified in [31], and with one exception ($E_8^{(1)}$ level 2) correspond to symmetries of the extended Dynkin diagram. $A_1^{(1)}$ has one non-trivial such symmetry, corresponding to the interchange of the nodes labeled '0' and '1'. This yields a permutation J of P_+^k given by $J(\lambda_0,\lambda_1) = (\lambda_1,\lambda_0)$, and corresponding to primary field $j = J0 = (0,k)$. Note that $Q_j(\lambda) = (-1)^{\lambda_1}$ and $R_j = k$. Thus $M(j)$ is a physical invariant whenever k is even. It is an automorphism invariant whenever $k \equiv 2 \pmod 4$.

The Galois action can be understood geometrically in terms of the affine Weyl group [25]. In general $\sigma(0) \neq 0$ and the signs ϵ_σ are both positive and negative. The S entries lie in the cyclotomic field $\mathbb{Q}(\xi_{f\,(k+h^\vee)})$, where f is a number depending only on the choice of algebra – e.g. $f = 4$ for A_1, and $f = 3,4,1,4,3$ for $X_r = E_6, E_7, E_8, F_4, G_2$, respectively (we will use this in the proof of Theorem 1 below). Also, $h^\vee = 12, 18, 30, 9, 4$ for those algebras.

A surprising feature is that almost all of the affine physical invariants can be understood in terms of the symmetries of the extended Dynkin diagram. For example, the classification for $A_1^{(1)}$ was done in [32]: apart from $M = M(0) = I$ (for all k) and $M = M(j)$ (for all even k), there are only 3 other physical invariants (at $k = 10, 16, 28$). The physical invariant classification for general affine algebras is still open; see [33] and references therein for a list of the algebras and levels for which it has been completed.

2.2 Finite group modular data

After this brief review of modular data in general RCFTs, and of affine modular data as specific examples, we turn to the modular data associated with a finite group G.

Fix a set R of representatives of each conjugacy class of G. So the identity e of G is in R, and more generally the centre $Z(G)$ of G is a subset of R.

For any $a \in G$, let K_a be the conjugacy class containing a. By $C_G(a)$ we mean the *centraliser* of a in G, i.e. the set of all elements in G which commute with a. $C_G(a)$ is a subgroup of G, and in fact $|G| = |K_a| \cdot |C_G(a)|$.

The primary fields of the G modular data are labelled by pairs (a,χ), where $a \in R$, and where χ is an irreducible character of $C_G(a)$. We will write $\Phi = \Phi(G)$ for the set of all these pairs. In this set, the vacuum '0' corresponds to $(e,1)$, with 1 the character of the trivial representation of $G = C_G(e)$. It will be convenient at times to identify (a,χ) with each $(g^{-1}ag, \chi^g)$, where $\chi^g(h) = \chi(ghg^{-1})$ is an irreducible character of $C_G(g^{-1}ag)$.

We set

$$S_{(a,\chi),(b,\chi')} = \frac{1}{|C_G(a)|\,|C_G(b)|}\sum_{g\in G(a,b)} \chi(gbg^{-1})^*\,\chi'(g^{-1}ag)^* \tag{2.12}$$

$$= \frac{1}{|G|}\sum_{g\in K_a,h\in K_b\cap C_G(g)} \chi(xhx^{-1})^*\,\chi'(ygy^{-1})^*, \tag{2.13}$$

$$T_{(a,\chi),(a',\chi')} = \delta_{a,a'}\delta_{\chi,\chi'}\frac{\chi(a)}{\chi(e)}, \tag{2.14}$$

where $G(a,b)=\{g\in G\,|\,agbg^{-1}=gbg^{-1}a\}$, and where x,y are any solutions to $g=x^{-1}ax$ and $h=y^{-1}by$. The equivalence of (2.12) and (2.13), and the fact that (2.13) does not depend on the choice of x,y, are easy arguments.

Note that the strange set $G(a,b)$ is precisely the set of all g for which $g^{-1}ag\in C_G(b)$ and $gbg^{-1}\in C_G(a)$. If there are no such g, then $G(a,b)$ is empty and the sum (and the matrix entry) would be equal to 0. A special case of this is Proposition 1(a) below.

S is symmetric and unitary, and gives rise (via Verlinde's formula) to non-negative integer fusion coefficients. The fusion coefficient $N_{(a,\chi_1),(b,\chi_2)}^{(c,\chi_3)}$ in fact has algebraic interpretations. For example, let ρ_i be representations corresponding to characters χ_i. Write $T(a)$ for a set of representatives of the left cosets of $C_G(a)$. Define the space

$$X=\bigoplus(\rho_1x)\otimes(\rho_2y), \tag{2.15}$$

where the sum is over all $x\in T(a)$, $y\in T(b)$ for which $(x^{-1}ax)\,(y^{-1}by)=c$, and where we interpret 'ρ_1x', 'ρ_2y' as belonging to the induced representations $(\rho_1)C_G(a)^G$, $(\rho_2)^G_{C_G(b)}$, respectively (we'll say more on induced representations below). X carries a representation of $C_G(c)$, using the usual coproduct. The fusion coefficient equals the multiplicity of ρ_3 in X [10]. (There are other interpretations of these fusion coefficients – see e.g. [16, 6, 7].)

Since a is in the centre of $C_G(a)$, $\chi(a)/\chi(e)$ in (2.14) will be an nth root of 1, where n is the order of a. Because $T_{00}=1$, the given normalisation of T corresponds to the central charge c being a multiple of 24; we are free to multiply (2.14) by any third root of 1, permitting c to be other multiples of 8.

In order to more fully exploit the formula (2.12), it is important to understand the notion of *induced character* (or representation) (see e.g. [3]). Given a representation $\rho:H\to W$ of a subgroup H of G, of character χ, we call a representation $\rho':G\to V$ of G an induced representation if V equals the direct sum of vector spaces W_a, for each coset $aH\in G/H$, where W_a is defined by $\rho'(aH)W=W_a$. An induced representation always exists and is unique up to isomorphism, and is denoted ρ_H^G. In terms of characters, we get the important formula (somewhat reminiscent of (2.12)):

$$\chi_H^G(g)=\frac{1}{|H|}\sum_{\substack{a\in G\\ a^{-1}ga\in H}}\chi(a^{-1}ga). \tag{2.16}$$

Suppose that $\chi'\in\mathrm{Irr}(C_G(b))$ is the restriction to $C_G(b)$ of some character $\overline{\chi}'$ defined on the group $G(a|b):=\langle G(a,b),C_G(b)\rangle$ (this happens fairly often in practice – see e.g. Ch.27 of

[1] for a relevant discussion). Then provided $G(a,b)\neq\emptyset$, (2.12) collapses to

$$S_{(a,\chi),(b,\chi')}=\frac{1}{|C_G(b)|}\,\overline{\chi}'(a)^*\,\chi^G_{C_G(a)}(b)^*. \tag{2.17}$$

Equation (2.17) for instance applies whenever χ' is identically 1. Another important instance of (2.17) is when z lies in the centre $Z(G)$, since then $C_G(z)=G$ and we get

$$S_{(a,\chi),(z,\chi')}=\frac{\chi(z)^*}{|C_G(a)|}\,\chi'(a)^*. \tag{2.18}$$

For instance, the quantum dimension of $(a,\chi)\in\Phi$ is

$$\frac{S_{(a,\chi),(e,1)}}{S_{(e,1),(e,1)}}=|K_a|\deg(\chi), \tag{2.19}$$

which amazingly enough is always a positive *integer* ! In fact from (2.16), (2.19) has a simple group-theoretic interpretation: it is the degree of the induced character $\chi^G_{C_G(a)}$. This integrality is very unusual, and shows that generically this finite group modular data is qualitatively different from affine data.

Equation (2.17) says we can expect many 0's in S. Any character $\chi\in\mathrm{Irr}(H)$ with degree $\chi(e)>1$ will have a zero – in fact $\chi(h)$ will be zero for at least $|Z(H)|\,(\chi(e)^2-1)$ group elements (see e.g. Ch. 23 of [1]). Hence the (z,χ) row will have 0's iff χ is not of degree 1. Dually, for any $b\in H$ there will be at least $k(H)-|C_H(b)|$ different $\chi'\in\mathrm{Irr}(H)$ with $\chi'(b)=0$, where $k(H)$ is the class number of H, so for instance the $(a,1)$ row will have 0's whenever e.g. there are elements in $C_G(a)$ whose centraliser in $C_G(a)$ is small. This seems to also be different from the affine case, where 0's are quite rare and fairly generically tend to be due to simple current fixed points.

Equation (2.18) also says

$$\frac{S_{(e,\chi),(a,\chi')}}{S_{(e,1),(a,\chi')}}=\chi(a)^*. \tag{2.20}$$

As long as we could identify by looking at S and T the primaries of the form $(e,\star)$, then the matrix S will contain the character table of G, and we would know that groups with different character tables would necessarily have different matrices S and T.

From (2.19), we see that the simple currents are precisely the pairs (z,φ), where z lies in the centre $Z(G)$ of G, and φ is a degree-1 character of G. Thus the group of simple currents is isomorphic to the direct product $Z(G)\times G/G'$, where G' is the commutator subgroup of G. The simple current $j=(z,\varphi)$ corresponds to permutation $J_{(z,\varphi)}(a,\chi)=(za,\varphi\chi)\in\Phi$, and charge $e^{2\pi\mathrm{i}Q}=\varphi(a)^*\chi(z)^*/\chi(1)$, which is always a root of unity as it should be. We get that $M(z,\varphi)$ (in the notation of the previous subsection) is always a physical invariant; it is an automorphism invariant iff z and φ have the same order.

Generic groups have many simple currents. A group for which $G=G'$ is called *perfect*; a group will have no non-trivial simple currents iff it is perfect and has trivial centre. For example this happens whenever G is non-cyclic simple. All perfect groups with small orders have been classified [34], and using this we can list all groups G with order $|G|<688128$ which have no non-trivial simple currents. The orders under 2000 of these groups are (see also Prop. 1(g)): 60, 168, 360, 504, 660, 960 (twice), 1092, 1344 (twice), and 1920.

When (and only when) G is abelian, all primaries will be simple currents, and hence the modular data will be rather trivial and uninteresting. This can also happen with the affine algebras: namely, the simply-laced algebras $A_r^{(1)}, D_r^{(1)}, E_6^{(1)}, E_7^{(1)}, E_8^{(1)}$ at level 1.

Charge conjugation takes (a,χ) to (a^{-1},χ^*), where the complex conjugate χ^* is the character of the contragredient representation. Now a^{-1} may not lie in our set R of conjugacy class representatives: recall that by (a^{-1},χ^*) we really mean $(g^{-1}a^{-1}g,\chi^{g\,*})\in\Phi$ where $g^{-1}a^{-1}g\in R$. This is not so trivial as it may seem: if a and a^{-1} are conjugate, (a,χ) and (a^{-1},χ^*) may be identified even though χ is complex-valued. An example is $\mathfrak{S}_3$, $a=(123)$, given below. Note though from (2.20) that the field $\mathbb{Q}(S)$ contains the field generated over $\mathbb{Q}$ by all character values $\chi(a)$, $\chi\in\mathrm{Irr}(G)$, $a\in G$, so if some characters of G are complex, the C won't be trivial.

The Galois symmetry is also straightforward. The character values of any group lie in the cyclotomic field $\mathbb{Q}(\xi_e)$ where $e=e(G)$ is the exponent of the group (the least common multiple of the orders of all group elements) – see Thm.8.7 of [3]. Hence the entries of S and T lie in the cyclotomic field $\mathbb{Q}(\xi_e)$, so the relevant Galois group is the multiplicative group $\mathbb{Z}_e^\times$. The Galois automorphism σ_ℓ takes (a,χ) to $(a^\ell,\sigma_\ell\circ\chi)$: $\sigma_\ell\circ\chi$ is the function obtained by applying σ_ℓ to each complex number $\chi(a)$; it is an irreducible character (of degree equal to that of χ) iff χ is [3]. Note another curiousity here: the Galois parities ϵ_ℓ are all identically equal to 1 ! This is very different from the generic affine situation. Note also that every Galois permutation fixes the vacuum $(e,1)$ (since every quantum dimension is rational), so large numbers of automorphism invariants will arise generically: whenever $\ell^2\equiv 1 \pmod{e(G)}$, the permutation σ_ℓ on Φ will define an automorphism invariant. We will return to this in §4.1.

Let us collect a few of the observations we have made here.

Proposition 1. (a) Choose any $(a,\chi),(a',\chi')\in\Phi$. If the order of a does not divide the exponent of $C_G(a')$, or if the order of a' does not divide the exponent of $C_G(a)$, then $S_{(a,\chi),(a',\chi')}=0$.

(b) The order of T *equals* (and not merely divides) the exponent of G.

(c) If $S(G)=S(H)$, then $|G|=|H|$.

(d) If $a^{-1}\notin K_a$, then the charge conjugate $C(a,\chi)\neq(a,\chi)$ for all $\chi\in\mathrm{Irr}(C_G(a))$. If $\chi\in\mathrm{Irr}(G)$ is not real-valued, then $C(z,\chi)\neq(z,\chi)$ for all $z\in Z(G)$.

(e) The quantum dimensions $S_{(a,\chi),(e,1)}/S_{(e,1),(e,1)}$ are always integers.

(f) The Galois parities $\epsilon_\sigma(a,\chi)$ are always $+1$, and the vacuum $(e,1)$ is fixed by all σ.

(g) The groups $G\neq\{e\}$ with at most 75 primaries, which have no non-trivial simple currents, are the simple groups $G=\mathfrak{A}_5,\mathrm{PSL}_2(\mathbb{F}_7),\mathfrak{A}_6,\mathrm{SL}_2(\mathbb{F}_8),\mathrm{PSL}_2(\mathbb{F}_{11})$, and $\mathfrak{A}_7$, with $(|G|,|\Phi|)=(60,22),(168,35),(360,44),(504,74),(660,58)$, and (2520,74).

(h) For any group $G\neq\{e\}$, there will be at least 5 physical invariants. If $G/Z(G)$ is nonabelian, then there will be at least $a+(b+c)^2$ physical invariants, where $a=|Z(G)|\,|G/G'|$, b equals the number of subgroups of $Z(G)$, and c equals the number of subgroups of G/G'.

The proof for (b) is the following. Choose any group H, and any $a \in Z(H)$ with prime-power order p^m. Then any irreducible representation of H will send a to the multiple of the identity matrix by some p^mth root of 1, i.e. $\chi(a)/\chi(e)$ will be a p^mth root of 1 for all $\chi \in \mathrm{Irr}(H)$. Now, that root of 1 must be primitive for some χ, as otherwise $a^{p^{m-1}}$ and e would have identical character values. Applying that fact to $H = C_G(a)$ gives us the desired order of T.

The proof of (c) is the comparison of $S_{(e,1),(e,1)}$ for G and H.

A constructive proof of (h) is given in §4.2. Incidentally, this bound can probably be significantly improved. $G = \mathbb{Z}_2$, with 6 physical invariants, is probably the lowest number. By comparison, affine algebras have relatively few physical invariants: e.g. both $B_\ell^{(1)}$ and $C_\ell^{(1)}$ at generic levels are expected to have only 2. ■

We have already remarked on several occasions above, that the finite group modular data are very different from affine modular data. At this stage (we will have more to say when we come to general twisted finite group modular data), this statement can be substantiated by noting that the two sets have a very small intersection.

Theorem 1. Let S and T be the Kac-Peterson matrices corresponding to an affine algebra $X_r^{(1)}$ at some level $k \geq 1$ (where X_r is simple). Let G be a finite group with $S(G) = S$ and $T(G) = \varphi T$ for some third root φ of 1. Then either

(i) $(X_r, k) = (E_8, 1)$ and $G = \{e\}$, or

(ii) $(X_r, k) = (D_{8n}, 1)$, and $G = \mathbb{Z}_2$.

Sketch of proof Write $n = k + h^\vee$. Recall that there is a Galois automorphism for any ℓ coprime to fn, where f for the exceptional algebras was given in §2.1. A consequence of the affine Weyl interpretation of the affine Galois permutation σ_ℓ of P_+^k is that the vectors $\sigma_\ell(\lambda + \rho)$ and $\ell\,(\lambda + \rho)$ have the same norm mod $2n$.

A nice way to handle the exceptional algebras is to check that $\ell\rho$ and ρ have the same norm (mod $2n$) for any ℓ coming from Galois (since all Galois automorphisms here will fix the vacuum 0). For instance, for E_7 we get $\rho^2 = 399/2$, so we see that $2n$ must divide $(\ell^2 - 1)\,399/2$. Now, the "Definition of 24" says that $\ell^2 - 1$ here can be replaced with 24: more precisely, the gcd of all numbers $\ell^2 - 1$, for ℓ coprime to fn, will equal $\gcd(24, (fn)^\infty)$. Hence, n must divide $2 \cdot 3^2 \cdot 7 \cdot 19$.

Now, $n > h^\vee = 18$. Also, if n is too big (i.e. if there is an ℓ coprime to fn such that $(h^\vee - 1)\ell < n$), then $\sigma_\ell(0)$ will equal $(m, \ell - 1, \ell - 1, \ldots, \ell - 1)$, where $m = n - 1 - \ell\,(h^\vee - 1)$, violating the result that 0 is fixed by all Galois automorphisms. For E_7, this is the condition that no ℓ, $1 < \ell < \frac{n}{17}$, can be coprime to $2n$.

We can now write down the possibilities for n: they are $19, 21, 38, 42, 57, 63$. Subtracting 18 gives the possible levels. $n = 19$ fails, since it would have to correspond to an abelian group with order $\frac{1}{S_{00}} = \sqrt{2}$. The remaining 5 possibilities all have non-integral quantum dimensions.

The other exceptional algebras are all done similarly. The only surviving n for G_2 are $n = 7, 8, 14$; for F_4 are $n = 12, 13, 18, 36, 39$; for E_6 are $n = 18, 24, 26, 36, 52$; and for E_8 are $n = 40, 48, 60, 62, 80, 120$. In all these cases, a weight with non-integral quantum dimension is easily found.

The best way to handle the classical algebras is to compute the quantum dimension of any weight in the Galois orbit of 0, and show it must be larger than 1, for some σ_ℓ. For now consider $k>1$ for A_r and C_r, and $k>2$ for B_r and D_r. Up to a sign, the quantum dimension of $\sigma_\ell(0)$ for the algebras A_r, B_r, C_r, D_r is, respectively,

$$\prod_{a=1}^{r}\frac{\sin(\pi\ell a/n)^{r+1-a}}{\sin(\pi a/n)^{r+1-a}}, \tag{2.21}$$

$$\prod_{a=0}^{r-1}\frac{\sin(\pi\ell\,(a+1/2)/n)}{\sin(\pi\,(a+1/2)/n)}\prod_{b=1}^{2r-2}\frac{\sin(\pi\ell b/n)^{[\frac{2r-b}{2}]}}{\sin(\pi b/n)^{[\frac{2r-b}{2}]}}, \tag{2.22}$$

$$\prod_{a=1}^{r-1}\frac{\sin(\pi\ell a/n)^{r-a}\,\sin(\pi\ell\,(a-1/2)/n)^{r-a}}{\sin(\pi a/n)^{r-a}\,\sin(\pi\,(a-1/2)/n)^{r-a}}\prod_{b=r}^{2r-1}\frac{\sin(\pi\ell b/2n)}{\sin(\pi b/2n)}, \tag{2.23}$$

$$\prod_{a=1}^{r-1}\frac{\sin(\pi\ell a/n)^{[\frac{2r-a+1}{2}]}}{\sin(\pi a/n)^{[\frac{2r-a+1}{2}]}}\prod_{b=r}^{2r-3}\frac{\sin(\pi\ell b/n)^{[\frac{2r-b-1}{2}]}}{\sin(\pi b/n)^{[\frac{2r-b-1}{2}]}}, \tag{2.24}$$

where $[x]$ here denotes the greatest integer not more than x. The absolute value of each of these is quickly seen to be greater than 1 unless $\ell\equiv\pm1 \pmod n$ (for A_r or D_r) or $\ell\equiv\pm1 \pmod{2n}$ (for B_r and C_r). This exhausts the possible ℓ coprime to fn only if the Euler totient $\varphi(n)\leq 2$ (for A_r and D_r) or $\varphi(2n)\leq 2$ (for B_r and C_r). By definition $\varphi(m)$ is the number of h, $1\leq h\leq m$, coprime to m; it is less than 3 only for $m=1,2,3,4,6$. So only A_1 level 2, A_2 level 3, A_1 level 4, and A_3 level 2 survive, but their central charges aren't multiples of 8.

The series A_r, D_{odd} and D_{even} at $k=1$ possess only simple currents (so would have to correspond to an abelian group G), and have the fusion groups $\mathbb{Z}_{r+1}, \mathbb{Z}_4, \mathbb{Z}_2\times\mathbb{Z}_2$ respectively. Abelian G has fusion group $G\times G$, so that leaves only D_{even} and $G=\mathbb{Z}_2$. D_n level 1 has $c=n$, concluding the argument. B_r level 1 has only 2 primaries so is covered e.g. by Theorem 2 below. The modular data of C_r level 1 is identical with that of A_1 level r.

Level 2 for D_r can be handled by requiring $|G|=\frac{1}{S_{00}}=2\sqrt{2r}$ and the quantum dimension $\frac{S_{\Lambda_r,0}}{S_{00}}=\sqrt{r}$ to both be integers (see [33] for the necessary S entries). For B_r at level 2, first read off from [33] that $|G|=\frac{1}{S_{00}}=2\sqrt{2r+1}$, so $2r+1=s^2$ for some odd integer s. Now, $T_{\Lambda_1,\Lambda_1}=\mathrm{i}\,\exp[-\pi\mathrm{i}\,\frac{1}{2s^2}]$ is a primitive s^2-root of 1, but $T(G)$ will have order dividing $|G|=2s$, and so those matrices cannot be equal. ■

For which groups will the number of primaries be low? Consider the formula

$$|\Phi(G)|=\sum_{a\in R}k(C_G(a)), \tag{2.25}$$

where $k(H)$ is the class number of H, i.e. the number of conjugacy classes in, or irreducible representations of, H. Note that the smaller $k(G)$ is, the fewer summands there will be in (2.25), the larger each conjugacy class $|K_a|$ will tend to be, so the smaller the centralisers $|C_G(a)|=\frac{|G|}{|K_a|}$ will tend to be, and the smaller the $k(C_G(a))$ in (2.25) will tend to be. Thus, we should expect $|\Phi|$ to grow with $k(G)$. The groups with class number less than 13 are classified [35]. This allows all G with at most 77 primaries to be listed. We make this argument precise in the proof of Thm. 2 below.

When G is abelian, it has $|G|^2$ primaries – this is the extreme case. The number of primaries for the even dihedral groups $\mathfrak{D}_{2n}$, and the quaternion group $\mathfrak{Q}_{2n}$, are both $2n^2+14$, which grows like $|G|^2/8$. The odd dihedral groups $\mathfrak{D}_m$, m odd, have $\frac{m^2+7}{2}$ primaries (see below). By comparison (using the data given in Ch.20 of [3]), $\mathrm{SL}_2(\mathbb{F}_q)$ has q^2+8q+9 primaries for q any power of an odd prime, and q^2+q+2 primaries for q any power of 2 – in either case, that number grows like $|G|^{2/3}$. Our computations so far suggest the following rule of thumb: the more abelian the group is, the messier it behaves (i.e. the more its primaries, the more its physical invariants, etc), while the closer the group is to being non-abelian simple, the better behaved it will be.

From this point of view, an interesting measure of how complicated a group is relative to its size, is the ratio

$$\mathcal{N}(G) := \frac{\log|\Phi(G)|}{\log|G|}. \tag{2.26}$$

It ranges from 0 to 2, with 2 achieved iff G is abelian. How low can $\mathcal{N}(G)$ be ? Some small values are $\mathcal{N}(\mathfrak{A}_5) \approx .75$, $\mathcal{N}(\mathfrak{A}_7) \approx .55$ and $\mathcal{N}(M_{11}) \approx .49$. Sporadic simple groups like the Monster should have $\mathcal{N}$ very small.

Theorem 2. There are precisely 33 groups G with at most 50 primaries:

- the abelian groups $G = \mathbb{Z}_1, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_4, \mathbb{Z}_2\times\mathbb{Z}_2, \mathbb{Z}_5, \mathbb{Z}_6, \mathbb{Z}_7$, with precisely $|G|^2$ primaries;
- the symmetric and alternating groups $\mathfrak{S}_3, \mathfrak{A}_4, \mathfrak{S}_4, \mathfrak{A}_5, \mathfrak{S}_5, \mathfrak{A}_6$ with $(|G|, |\Phi|) = (6,8)$, (12,14),(24,21),(60,22),(120,39),(360,44);
- the (semi)dihedral and quaternion groups $\mathfrak{D}_5, \mathfrak{D}_4, \mathfrak{Q}_4, \mathfrak{D}_7, \mathfrak{D}_9, \mathfrak{D}_8, \mathfrak{Q}_8, S\mathfrak{D}_8$ with $(|G|, |\Phi|) = (10,16), (8,22), (8,22), (14,28), (18,43), (16,46), (16,46), (16,46)$;
- the Frobenius groups $\mathbb{Z}_5 \times_f \mathbb{Z}_4$, $\mathbb{Z}_7 \times_f \mathbb{Z}_3$, $\mathbb{Z}_3^2 \times_f \mathbb{Z}_2$, $\mathbb{Z}_3^2 \times_f \mathbb{Z}_4$, $\mathbb{Z}_3^2 \times_f \mathfrak{Q}_4$, $\mathbb{Z}_7 \times_f \mathbb{Z}_6$, $\mathbb{Z}_{11} \times_f \mathbb{Z}_5$ with $(|G|, |\Phi|) = (20,22), (21,32), (18,44), (36,36), (72,32), (42,44), (55,49)$;
- the remaining groups $\mathbb{Z}_2 \times \mathfrak{S}_3$, DC_3, $\mathrm{PSL}_2(\mathbb{F}_7)$, and $\mathrm{SL}_2(\mathbb{F}_3)$, with $(|G|, |\Phi|) = (12,44)$, (12,32), (168,35), (24,42).

The semidihedral group $S\mathfrak{D}_{4m}$ is defined by the presentation

$$S\mathfrak{D}_{4n} = \langle r, s \mid r^{4n} = s^2 = e,\, srsr = r^{2n}\rangle. \tag{2.27}$$

and has order $8m$. The other groups are defined in [35]. An interesting consequence of Theorem 2, as mentioned earlier, is that all groups with $|\Phi(G)| \leq 50$ are uniquely determined by their matrix S.

A useful quantity for proving Thm. 2 is $h(n)$, the minimum possible class number $k(H)$ for H with order n and non-trivial centre. Knowing $h(n)$ gives a lower bound for $|\Phi(G)|$ once we know the orders $|C_G(a)|$ of its centralisers. [35] can be used for the smaller n, basic results on the classification of finite groups, as well as the congruence $k(H) \equiv |H| \pmod{16}$ when $|H|$ is odd, give us other n. What we find is e.g. $h(n) = n$ for all $n \leq 27$, except for $n = 8, 12, 16, 18, 20, 24$ (with $h(n) = 5, 6, 7, 9, 8, 7$, resp). Also useful is the largest value $\ell(k)$ of $|H|$, for H with non-trivial centres and class number $k(H) \leq k$. For instance for $k = 1, 2, \ldots, 8$ we get $\ell = 1, 2, 3, 4, 8, 12, 24, 48$.

Consider the smallest $|\Phi|$ can be if $k(G) \geq 13$. Note that at most one $a \in R$ can have a centraliser $C_G(a)$ with class number 2 — otherwise if there were two then together exactly $|G|/\ell(2) + |G|/\ell(2) > |G| - 1$ elements would be in the (disjoint) conjugacy classes $K_a \cup K_b \subset G - e$. More generally, at most $\ell(k) - 1$ a's could have $k(C_G(a)) \leq k$. Thus (2.25) will be bounded below by $13 + 2 + 3 + 4 + 4 \cdot 5 + 4 \cdot 6 + 7 = 73$. Tightening the argument (e.g. $\frac{1}{2} + \frac{1}{3} + \frac{1}{4}$ is too big) gives $|\Phi| \geq 78$. Thus using the tables of [35] we would be able to find all groups with at most 77 primaries. Similarly, to do $|\Phi| \leq 50$ it is enough to consider $k(G) \leq 9$. In [35] are also given the orders of the centralisers $C_G(a)$. It is now straightforward to get the Theorem. ∎

3 Examples

In this section we give a number of explicit examples of untwisted modular data. We also identify their physical invariants in some cases, leading to a perplexing situation we will discuss more fully next section.

Incidentally, a simple construction is direct product: the modular data for the direct product $G \times H$ is easily obtained from that of G and H. For example, $\Phi(G \times H) = \Phi(G) \times \Phi(H)$, $S(G \times H)$ is the Kronecker matrix product $S(G) \otimes S(H)$, etc. Of course, semi-direct product in general will be much more difficult to work out.

3.1 Abelian groups

Abelian G (untwisted) is trivial to work out, but also very uninteresting. Write G in the following canonical way: $G \cong \mathbb{Z}_{d_1} \times \mathbb{Z}_{d_2} \times \cdots \times \mathbb{Z}_{d_s}$ where $d_1|d_2|\cdots|d_s$. For convenience, define a bilinear form on $\mathbb{Z}^s$ by $\langle m, n\rangle = \sum_i \frac{m_i n_i}{d_i}$. We can identify Φ here with the $2s$-tuple $(m, n) \in \mathbb{Z}^s \times \mathbb{Z}^s$, where $0 \leq m_i \leq d_i - 1$ and $0 \leq n_i \leq d_i - 1$: in particular, m corresponds to the group element $(m_1, m_2, \ldots, m_s) \in \mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_s}$, and n corresponds to the character φ_n of G defined by $\varphi_n(m) = \exp[2\pi \mathrm{i}\langle m, n\rangle]$. The matrices S and T are given by

$$S_{(m,n),(m',n')} = \frac{1}{|G|}\exp[-2\pi\mathrm{i}(\langle m', n\rangle + \langle m, n'\rangle)], \qquad T_{(m,n),(m,n)} = \exp[2\pi\mathrm{i}\langle m, n\rangle]. \tag{3.1}$$

All $(m, n) \in \Phi$ are simple currents, with composition given by pairwise addition. Charge conjugation takes (m, n) to $(-m, -n)$, and more generally the ℓth Galois automorphism (for $\ell \in \mathbb{Z}_{d_s}^{\times}$) sends (m, n) to $(\ell m, \ell n)$.

G will have many physical invariants, but they can all be most elegantly interpreted using lattices as was explained in [36] (the standard reference for lattice theory is [37]). In particular let Λ be the $4s$-dimensional integral indefinite lattice, given by the orthogonal direct sum $\Lambda = \sqrt{d_1}II_{2,2} \oplus \sqrt{d_2}II_{2,2} \oplus \cdots \oplus \sqrt{d_s}II_{2,2}$, where $II_{2,2} = II_{1,1} \oplus II_{1,1}$ is the unique 4-dimensional even self-dual indefinite lattice. Then there is a natural one-to-one bijection between the physical invariants of G and the even self-dual 'gluings' of Λ, i.e. the even self-dual $4s$-dimensional lattices containing Λ. For instance, for $G = \mathbb{Z}_p$ (p prime), one finds that there are precisely 6 (if $p = 2$) or 8 (if $p > 2$) physical invariants. Two of these are

$$(\sum_{i=0}^{p-1} \mathrm{ch}_{i0})\,(\sum_{j=0}^{p-1} \mathrm{ch}^*_{0j}) \qquad \text{and} \qquad \sum_{i,j=0}^{p-1} \mathrm{ch}_{ij}\, \mathrm{ch}^*_{-j,-i}\,, \tag{3.2}$$

using obvious notation.

The non-abelian groups are much more interesting, and we turn to them in the next subsection. The number of non-abelian groups of order $n \leq 50$ are 1 (for $n = 6, 10, 14, 21, 22, 26, 34, 38, 39, 46$), 2 ($n = 8, 27, 28, 44$), 3 ($n = 12, 18, 20, 30, 50$), 5 (for $n = 42$), 9 (for $n = 16$), 10 (for $n = 36$), 11 (for $n = 40$), 12 (for $n = 24$), 44 (for $n = 32$), 47 (for $n = 48$), and 0 otherwise.

3.2 Some infinite series

It is not hard to work out S and T for the infinite series $\mathfrak{D}_n$ (dihedral) and $\mathfrak{Q}_{2n}$ (quaternion):

$$\mathfrak{D}_n = \langle r, s \mid r^n = s^2 = e,\, rsrs = e\rangle, \tag{3.3}$$

$$\mathfrak{Q}_{2n} = \langle r, s \mid r^{2n} = e,\, s^2 = r^n,\, rsrs^{-1} = e\rangle, \tag{3.4}$$

The character tables of $\mathfrak{D}_{4n}$ and $\mathfrak{Q}_{4n}$ are identical.

Consider first the even dihedral groups $\mathfrak{D}_{2n}$, of order $4n$. $\mathfrak{D}_{2n}$ has $n + 3$ conjugacy classes, with representatives $R = \{e,\, r^n,\, r^k\ (1 \leq k \leq n-1),\, s,\, sr\}$. It has an equal number of irreducible representations; 4 are one-dimensional and $n - 1$ are two-dimensional. The characters and centralisers of the various classes are indicated in the following table.

	e	r^n	r^k	s	sr
ψ_0	1	1	1	1	1
ψ_1	1	1	1	-1	-1
ψ_2	1	$(-1)^n$	$(-1)^k$	1	-1
ψ_3	1	$(-1)^n$	$(-1)^k$	-1	1
χ_i	2	$2(-1)^i$	$2\cos\frac{\pi i k}{n}$	0	0
$C_G(g)$	$\mathfrak{D}_{2n}$	$\mathfrak{D}_{2n}$	$\mathbb{Z}_{2n}$	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_2$

It follows that the number of primaries is equal to $|\Phi| = 2n^2 + 14$. Denote by ψ_i $(0 \leq i < 2n)$ the characters of $\langle r\rangle \cong \mathbb{Z}_{2n}$, given by $\psi_i(r^k) = \xi_{2n}^{ik}$. Likewise, denote by ϕ_{ab} and φ_{ab} $(0 \leq a, b \leq 1)$ the characters of $\langle s, r^n\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ and $\langle rs, r^n\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, respectively, where $\phi_{ab}(s^k r^{n\ell}) = \varphi_{ab}((rs)^k r^{n\ell}) = (-1)^{ak+b\ell}$.

The values of the diagonal entries of T follow directly from (2.14). Eq. (2.18) easily computes any S entry involving e or r^n. The remaining non-zero S entries are

$$S_{(r^k,\psi_i),(r^\ell,\psi_j)} = \frac{1}{n}\cos(\pi\frac{\ell i + kj}{n}), \tag{3.5}$$

$$S_{(s,\phi_{ab}),(s,\phi_{cd})} = S_{(rs,\varphi_{ab}),(rs,\varphi_{cd})} = \frac{1}{4}\cdot\begin{cases} (-1)^{a+c} + (-1)^{a+b+c+d} & \text{if } n \text{ even,} \\ (-1)^{a+c} & \text{if } n \text{ odd,} \end{cases} \tag{3.6}$$

and in addition for n odd $S_{(s,\phi_{ab}),(sr,\varphi_{cd})} = \frac{1}{4}(-1)^{a+b+c+d}$.

The odd dihedral groups $\mathfrak{D}_{2n+1}$ can be worked out in the same way. $\mathfrak{D}_{2n+1}$ has $n+2$ conjugacy classes, with representatives $R = \{e,\, r^k\ (1 \le k \le n),\, s\}$. It has two one-dimensional representations, and n two-dimensional representations. The characters and centralisers of the various classes are reproduced in the following table.

	e	r^k	s
ψ_0	1	1	1
ψ_1	1	1	-1
χ_i	2	$2\cos\frac{2\pi ik}{2n+1}$	0
$C_G(g)$	$\mathfrak{D}_{2n+1}$	$\mathbb{Z}_{2n+1}$	$\mathbb{Z}_2$

One finds that the number of primary fields is $|\Phi| = 2n^2 + 2n + 4$. Write ψ_i for the characters of $\langle r\rangle \cong \mathbb{Z}_{2n+1}$ as before, and φ_i for the obvious two characters of $\langle s\rangle \cong \mathbb{Z}_2$. The calculation of S and T proceeds like for the even dihedral groups:

$$S_{(r^k,\psi_i),(r^\ell,\psi_j)} = \frac{2}{2n+1}\cos(2\pi\frac{kj+\ell i}{2n+1}), \tag{3.7}$$

$$S_{(s,\varphi_i),(s,\varphi_j)} = \frac{1}{2}(-1)^{i+j}. \tag{3.8}$$

The semidihedral groups $S\mathfrak{D}_{4m}$ (2.27) have the same matrix T as $\mathfrak{D}_{4m}$, but their matrix S is always complex.

Next turn to the quaternions $\mathfrak{Q}_{2n}$. It has $n+3$ conjugacy classes, with representatives $R = \{e,\, r^n,\, r^k\ (1 \le k \le n-1),\, s,\, sr\}$. It has 4 one-dimensional and $n-1$ two-dimensional representations. The characters and centralisers of the various classes are indicated in the following table (put $\iota = 1$ for n even, and $\iota = \mathrm{i}$ for n odd).

	e	r^n	r^k	s	sr
ψ_0	1	1	1	1	1
ψ_1	1	1	1	-1	-1
ψ_2	1	$(-1)^n$	$(-1)^k$	ι	$-\iota$
ψ_3	1	$(-1)^n$	$(-1)^k$	$-\iota$	ι
χ_i	2	$2(-1)^i$	$2\cos\frac{\pi ik}{n}$	0	0
$C_G(g)$	$\mathfrak{Q}_{2n}$	$\mathfrak{Q}_{2n}$	$\mathbb{Z}_{2n}$	$\mathbb{Z}_4$	$\mathbb{Z}_4$

The number of primaries is equal to $|\Phi| = 2n^2 + 14$. Denote by ψ_i the $2n$ characters of $\langle r\rangle \cong \mathbb{Z}_{2n}$, and by ϕ_a and φ_a the 4 characters of $\langle s\rangle \cong \mathbb{Z}_4$ and $\langle rs\rangle \cong \mathbb{Z}_4$, respectively. The nonzero S entries not involving e or r^n are

$$S_{(r^k,\psi_i),(r^\ell,\psi_j)} = \frac{1}{n}\cos(\pi\frac{\ell i + kj}{n}), \tag{3.9}$$

$$S_{(s,\phi_a),(s,\phi_b)} = S_{(rs,\varphi_a),(rs,\varphi_b)} = \frac{1}{4}\cdot\begin{cases} \mathrm{i}^{a-b} & \text{if } n \text{ odd,} \\ 2\cos(\pi\frac{a+b}{2}) & \text{if } n \text{ even,}\end{cases} \tag{3.10}$$

and in addition for n odd $S_{(s,\phi_a),(sr,\varphi_b)} = \frac{1}{4}\mathrm{i}^{a+b}$.

For our final example in this subsection, we will consider the series of non-abelian simple groups, $\mathrm{SL}_2(\mathbb{F}_q)$ for $q = 2^n$. Note that $\mathrm{SL}_2(\mathbb{F}_2) \cong \mathfrak{S}_3$ and $\mathrm{SL}_2(\mathbb{F}_4) \cong \mathfrak{A}_5$.

The order of $\mathrm{SL}_2(\mathbb{F}_q)$ is $q(q^2-1)$. There are $q+1$ conjugacy classes whose representatives can be chosen in

$$R = \left\{e,\, \iota = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},\, \alpha^a = \begin{pmatrix} s^a & 0 \\ 0 & s^{-a} \end{pmatrix},\, \beta^b \;:\; 1 \le a \le \tfrac{q-2}{2} \text{ and } 1 \le b \le \tfrac{q}{2}\right\}, \tag{3.11}$$

where s is any generator of the cyclic multiplicative group $\mathbb{F}_q^*$, and where β is an element of order $q+1$ (its exact form is not important). The characters (the labels i and j run from 1 to $\frac{q}{2}-1$ and $\frac{q}{2}$ respectively) and centralisers are given below, from which one finds the number of primaries equals $|\Phi| = q^2+q+2$, as mentioned before.

	e	ι	α^a	β^b
ψ_0	1	1	1	1
ψ_1	q	0	1	-1
χ_i	$q+1$	1	$2\cos\frac{2\pi ia}{q-1}$	0
θ_j	$q-1$	-1	0	$-2\cos\frac{2\pi jb}{q+1}$
$C_G(g)$	$\mathrm{SL}_2(\mathbb{F}_q)$	$\mathbb{Z}_2^n$	$\mathbb{Z}_{q-1}$	$\mathbb{Z}_{q+1}$

Because $2, q-1, q+1$ are pairwise coprime, Proposition 1(a) says that the only potentially nonvanishing $S_{(a,\chi),(b,\phi)}$ entries have $a = e$ or $b = e$, or $a = b = \iota$, or a, b are both powers of α, or a, b are both powers of β. The relevant sets $G(g,h)$ here are $G(\iota,\iota) = \cup_{a=0}^{q-2}\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)\alpha^a$, $G(\alpha^a,\alpha^b) = \langle a\rangle \cup \langle a\rangle\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and $G(\beta^a,\beta^b) = \langle\beta\rangle \cup \langle\beta\rangle\gamma$ where $\gamma^{-1}\beta\gamma = \beta^{-1}$.

Explicitly writing down its matrix S would require the evaluation of some interesting character sums, something we have not yet done.

3.3 The physical invariants for $\mathfrak{S}_3$

The non-abelian group of smallest order is $\mathfrak{S}_3$. Its character table is

	e	(123)	(12)
ψ_0	1	1	1
ψ_1	1	1	-1
ψ_2	2	-1	0
$C_G(g)$	$\mathfrak{S}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_2$

The modular data for $\mathfrak{S}_3$ will have 8 primary fields: (e, ψ_i) for $i = 0, 1, 2$; $((123), \varphi_k)$, $k = 0, 1, 2$, for the 3 characters $a \mapsto \xi_3^{ak}$ of $\mathbb{Z}_3$; and $((12), \varphi'_k)$, $k = 0, 1$, for the 2 characters $b \mapsto (-1)^{bk}$ of $\mathbb{Z}_2$. For convenience label these primaries $0, 1, \ldots, 7$.

Since $\mathfrak{S}_3 \cong \mathfrak{D}_3$, we can read $S(\mathfrak{S}_3)$ and $T(\mathfrak{S}_3)$ off from the previous subsection:

$$S = \frac{1}{6}\begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 \\ 1 & 1 & 2 & 2 & 2 & 2 & -3 & -3 \\ 2 & 2 & 4 & -2 & -2 & -2 & 0 & 0 \\ 2 & 2 & -2 & 4 & -2 & -2 & 0 & 0 \\ 2 & 2 & -2 & -2 & -2 & 4 & 0 & 0 \\ 2 & 2 & -2 & -2 & 4 & -2 & 0 & 0 \\ 3 & -3 & 0 & 0 & 0 & 0 & 3 & -3 \\ 3 & -3 & 0 & 0 & 0 & 0 & -3 & 3 \end{pmatrix}, \tag{3.12}$$

$$T = \mathrm{diag}(1, 1, 1, 1, \xi_3, \xi_3^2, 1, -1). \tag{3.13}$$

There is one non-trivial simple current (e, ψ_1) (namely primary #1), identifiable by the 1 in the corresponding entry of the 0th row of S. Since all entries of S are rational, the charge conjugation and the other Galois permutations σ_ℓ are trivial. Incidentally, groups G for which $S(G)$ is rational are rare; that property requires that the exponent of G divides 24 (e.g. the exponent of $\mathfrak{S}_3$ is 6). To see that, apply Prop.1(b) to the equation $(T_{a,a})^{\ell^2} = T_{\sigma_\ell a, \sigma_\ell a}$.

There are precisely 32 physical invariants for $\mathfrak{S}_3$. Write ch_i for the CFT character corresponding to the ith primary. The automorphism invariants are $M = I$, and the one (call it M') which interchanges $2 \leftrightarrow 3$ and fixes everything else. Extending by the simple current gives us 3 invariants:

$$|\mathrm{ch}_0 + \mathrm{ch}_1|^2 + 2|\mathrm{ch}_2|^2 + 2|\mathrm{ch}_3|^2 + 2|\mathrm{ch}_4|^2 + 2|\mathrm{ch}_5|^2 + k(\mathrm{ch}_2\mathrm{ch}_3^* + \mathrm{ch}_3\mathrm{ch}_2^* - |\mathrm{ch}_2|^2 - |\mathrm{ch}_3|^2), \tag{3.14}$$

for $k = 0, 1, 2$. Write $s_1 := \mathrm{ch}_0 + \mathrm{ch}_1 + \mathrm{ch}_2 + \mathrm{ch}_3$, $s_2 := \mathrm{ch}_0 + \mathrm{ch}_1 + 2\mathrm{ch}_2$, $s_3 := \mathrm{ch}_0 + \mathrm{ch}_1 + 2\mathrm{ch}_3$, $s_4 := \mathrm{ch}_0 + \mathrm{ch}_2 + \mathrm{ch}_6$, and $s_5 := \mathrm{ch}_0 + \mathrm{ch}_3 + \mathrm{ch}_6$; then $s_i s_j^*$ is a physical invariant. The final ones are

$$|\mathrm{ch}_0 + \mathrm{ch}_1 + \mathrm{ch}_2 + \mathrm{ch}_3|^2 + k(\mathrm{ch}_2\mathrm{ch}_3^* + \mathrm{ch}_3\mathrm{ch}_2^* - |\mathrm{ch}_2|^2 - |\mathrm{ch}_3|^2), \tag{3.15}$$

for $k = \pm 1$.

In the next section we manage to identify general constructions yielding most of these physical invariants. Two however remain unexplained: the automorphism invariant M', and the modular invariance of the sum s_1.

Note that in the basis defined by ($\xi = \xi_3 = \exp[2\pi \mathrm{i}/3]$)

$$\begin{aligned}(e_0, \ldots, e_7) \;=\; (&\mathrm{ch}_0 - \mathrm{ch}_1\,,\, \mathrm{ch}_0 + \mathrm{ch}_1 + 2\mathrm{ch}_2\,,\, \mathrm{ch}_0 + \mathrm{ch}_1 - \mathrm{ch}_2\,,\, \mathrm{ch}_3 + \mathrm{ch}_4 + \mathrm{ch}_5\,,\\ &\mathrm{ch}_3 + \xi\mathrm{ch}_4 + \xi^2\mathrm{ch}_5\,,\, \mathrm{ch}_3 + \xi^2\mathrm{ch}_4 + \xi\mathrm{ch}_5\,,\, \mathrm{ch}_6 + \mathrm{ch}_7\,,\, \mathrm{ch}_6 - \mathrm{ch}_7), \end{aligned} \tag{3.16}$$

the matrices S and T become permutation matrices (both non diagonal), corresponding respectively to the permutations (06)(23)(45) and (345)(67) of $\mathfrak{S}_8$ (written in terms of cycles). From this, it is not difficult to compute the dimension (over $\mathbb{C}$) of the commutant of S and T, which turns out to be 11. Hence the 32 physical invariants are not linearly independent (e.g. $s_2 + s_3 = 2s_1$ and $2s_4 - 2s_5 = s_2 - s_3$).

Incidentally, it is a consequence of [6] that for any G, such a basis can always be found in which S and T are permutation matrices. In particular, the characters $\mathrm{ch}_{(a,\chi)}$, for $(a,\chi) \in \Phi$, can be thought of as a function on $G \times G$ taking any pair (xax^{-1}, xbx^{-1}) to $\chi(b)$ when a and b commute, and sending all other pairs in $G \times G$ to 0. They span a space called $C^0(G_{comm})$ in [12]. Now choose any commuting pair $(a,b) \in G \times G$, and define the function $f_{(a,b)}$ to be identically 1 on the set (xax^{-1}, xbx^{-1}) and 0 elsewhere. Then these functions form another basis for $C^0(G_{comm})$. For this choice of basis, our representation of $\mathrm{SL}_2(\mathbb{Z})$ becomes manifestly a permutation representation: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ acts on the right by sending $f_{(g,h)}$ to $f_{(g^a h^c, g^b h^d)}$. For instance, we recover T by noting that

$$\mathrm{ch}_{(a,\chi)}(a, ab) = \chi(ab) = \frac{\chi(a)}{\chi(e)}\,\chi(b) = T_{(a,\chi),(a,\chi)}\mathrm{ch}_{(a,\chi)}\ . \tag{3.17}$$

(See equation (5.12) of [12] for the corresponding matrix S calculation.) It is a very special property of the (untwisted) finite group modular data that this $\mathrm{SL}_2(\mathbb{Z})$ representation is actually a permutation representation. For general modular data, usually the matrix T won't even be conjugate to a permutation matrix. Two examples are the simplest affine data, i.e. $A_1^{(1)}$ at level 1, and the simplest twisted group data, $G = \mathbb{Z}_2$ with twist α_1 (see §6.1), whose matrices T are respectively

$$\mathrm{diag}\,(\exp[-\pi\mathrm{i}/12], \exp[5\pi\mathrm{i}/12]) \qquad \text{and} \qquad \mathrm{diag}\,(1, 1, \mathrm{i}, -\mathrm{i})\ . \tag{3.18}$$

The first cannot be similar to a permutation matrix because it does not have 1 as an eigenvalue, while the second cannot because an order 4 permutation matrix has -1 as an eigenvalue.

Incidentally, the analogue of this permutation representation for the twisted finite group data is known [11], and will be discussed in §5.3.

A total of 32 physical invariants for such a small number of primaries is completely unprecedented from the more familiar WZW situation. The orthogonal algebras at level 2 are the worst behaved affine cases, but e.g. $D_{16}^{(1)}$ at level 2 has 23 primaries but only 22 physical invariants. But $\mathfrak{S}_3$, being a dihedral group, is nearly abelian so should be especially bad in this respect.

It seems that the twisted data is better behaved in this respect than the untwisted. For example, we find, combining the discussion of §6.3 with the results of [33], that the twisted modular data for $\mathfrak{S}_3$ has only 9 physical invariants.

4 Making sense of all the physical invariants

We learned in the previous section that there is a surprising number of physical invariants associated to finite group modular data. In this section we try to tame the zoo!

The conventional wisdom in RCFT is that all physical invariants are constructed in two ways: as extensions of chiral algebras, and as automorphisms of those chiral algebras (called automorphism invariants in the unextended case). In the affine case, the physical invariants are almost always 'obvious' in hindsight, as symmetries of the appropriate Dynkin diagram (e.g. simple currents and charge conjugation) are directly responsible for almost all WZW physical invariants. It would be nice to find the analogous statement for the finite group modular data.

A rich supply of physical invariants always comes from simple currents [26, 27]. As these are well-understood, we won't say any more than we did in §2.

4.1 Automorphism invariants

The outer automorphism group $\mathrm{Out}(G)$ provides a systematic way of constructing automorphism invariants. Choose any $\pi \in \mathrm{Out}(G)$. Then π induces an invertible homomorphism $C_G(a) \to C_G(\pi a)$ between centralisers. Define χ^π by the formula $\chi^\pi(g) = \chi(\pi^{-1}g)$ — if $\chi \in \mathrm{Irr}(C_G(a))$ then $\chi^\pi \in \mathrm{Irr}(C_G(\pi a))$. Consider the permutation of Φ defined by $(a, \chi) \mapsto (\pi a, \chi^\pi)$. Then it commutes with S and T and hence defines an automorphism invariant.

Many (but not all) outer automorphisms can be interpreted in the following way: Any time a given group G is a normal subgroup of another group $\widehat{G}$, then any element $\hat{g} \in \widehat{G}$ defines an automorphism of G by conjugation: $g \mapsto \hat{g}g\hat{g}^{-1}$.

Examples of these groups are $\mathrm{Out}(\mathbb{Z}_n) \cong \mathbb{Z}_n^\times$; $\mathrm{Out}(\mathfrak{D}_{2k+1}) \cong \mathbb{Z}_{2k+1}^\times/\{\pm 1\}$; for $n \geq 3$, $\mathrm{Out}(\mathfrak{S}_n) = \{1\}$ and $\mathrm{Out}(\mathfrak{A}_n) \cong \mathbb{Z}_2$, except $\mathrm{Out}(\mathfrak{S}_6) \cong \mathbb{Z}_2$ and $\mathrm{Out}(\mathfrak{A}_6) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Another systematic source are the Galois automorphism invariants — there is one of these for every $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(S)/\mathbb{Q})$ with $\ell^2 \equiv 1$ modulo the exponent $e(G)$ of G, and the permutation is simply given by the Galois permutation, which we may write $\sigma_\ell(a, \chi) = (a^\ell, \sigma_\ell \chi)$. If there are s distinct primes which divide $e(G)$, then there will be precisely

$$\begin{cases} 2^{s-1} & \text{if } e(G) \equiv 2 \pmod 4 \\ 2^{s+1} & \text{if } e(G) \equiv 0 \pmod 8 \\ 2^s & \text{otherwise} \end{cases} \tag{4.1}$$

such ℓ, and hence that number of Galois automorphism invariants (though these won't necessarily be distinct, if $\mathbb{Q}(S)$ is smaller than $\mathbb{Q}(\xi_{e(G)})$).

A final source of automorphism invariants is *discrete torsion* [19]. This involves the language of cohomology — see §5.1 for the appropriate definitions. Take any 2-cocycle $\beta \in Z^2(G, U(1))$ for G. This β is entirely independent of the 2-cocycles β_a we discuss in §5.2. For each conjugacy class representative $a \in R$, define $e_a(b) := \beta(a, b)\beta(b, a)^{-1}$. This will be a 1-dimensional representation of $C_G(a)$. Define the permutation of primary fields $(a, \chi) \in \Phi$ sending (a, χ) to $(a, e_a\chi)$. It is easy to check that this commutes with S and T and thus defines an automorphism invariant.

When is this permutation trivial ? Iff each $e_a(b) = 1$, for all $b \in C_G(a)$. In other words, iff each $a \in G$ is β-regular (see §5.1 for the definition). One thing this means is that cohomologous β give the same discrete torsion. So the possibilities for discrete torsion are given by the classes in the finite abelian group $H^2(G, U(1)) = M(G)$, known as the Schur multiplier of G. Each class in $M(G)$ will usually but not always give rise to a different automorphism invariant. Incidentally, this construction applies to untwisted, CT twisted, or non-CT twisted, modular data (see next section for the twisted modular data).

So discrete torsion will always give $|M(G)|$ automorphism invariants, although these are not necessarily all distinct. They are all distinct for $\mathbb{Z}_n^2$, $\mathbb{Z}_n^3$, and $\mathfrak{D}_{even}$.

4.2 Chiral extensions

Making sense of the large number of physical invariants also requires finding the "generic" chiral extensions. Many will come from simple currents. It is tempting to suspect that another rich source could be from certain normal subgroups $N \triangleleft G$.

In particular, Clifford theory (which is concerned with the theory of induced and restricted characters) can be used to explicitly relate the modular data of a group G to that of its normal subgroups H. The formulas we have obtained however are complicated enough that at this point we have been unable to determine explicit relations between the physical invariants of G and H. Nevertheless, the situation is sufficiently analogous to that of conformal embeddings [43] in affine algebras that we expect this to likewise be a rich source of chiral extensions. Incidentally, there are also relations between the modular data of G and G/N.

Nevertheless we have found some interesting generic chiral extensions. Choose any central elements $z, z' \in Z(G)$ and any degree-1 characters $\varphi, \psi \in \mathrm{Irr}(G)$. Define the combinations

$$s_{z,z'} := \sum\nolimits_{\chi \in \mathrm{Irr}(G)} \chi(z') \, \mathrm{ch}_{(z,\chi)} \tag{4.2}$$

$$s'_{\varphi,\psi} := \sum\nolimits_{g \in R} \psi(g) \, \mathrm{ch}_{(g,\varphi)} \tag{4.3}$$

Then it is easy to verify, using (2.17), Frobenius reciprocity, and the orthogonality relations for characters, that $s_{z,z'}(-1/\tau) = s_{z',z^{-1}}(\tau)$ and $s'_{\varphi,\psi}(-1/\tau) = s'_{\psi,\varphi^*}(\tau)$, while $s_{z,z'}(\tau+1) = s_{z,zz'}(\tau)$ and $s'_{\varphi,\psi}(\tau+1) = s'_{\varphi,\varphi\psi}(\tau)$. To see this, consider the coefficient of $\mathrm{ch}_{(a,\chi)}(\tau)$ in $s'_{\varphi,\psi}(-1/\tau)$: it will be

$$\sum_{g \in R} \psi(g) \, \frac{1}{|G|} \, \varphi(a)^* \chi^G_{C_G(a)}(g)^* \tag{4.4}$$

which is $\varphi(a)^*$ times the coefficient of ψ in $\chi^G_{C_G(a)}$, i.e. $\varphi(a)^*$ times the coefficient of χ in $\psi|_{C_G(a)}$, i.e. $\varphi(a)^* \delta_{\psi,\chi}$.

Choose any subgroups $A \leq Z(G)$, $B \leq G/G'$, and define the sums

$$s(A) = \frac{1}{|A|} \sum_{z,z' \in A} s_{z,z'} = \sum_{z \in A} \sum_{\chi : \ker(\chi) \geq A} \chi(e) \, \mathrm{ch}_{z,\chi} \tag{4.5}$$

$$s'(B) = \frac{1}{|B|} \sum_{\varphi,\psi \in B} s'_{\varphi,\psi} = \sum_{\psi \in B} \sum_{g \in \ker(B) \cap R} \mathrm{ch}_{g,\psi} \tag{4.6}$$

By $\ker(\chi)$ we mean all $g \in G$ for which $\chi(g) = \chi(e)$, and by $\ker(B)$ we mean the intersection of all $\ker(\psi)$, where $\psi \in B$ is identified with a degree-1 character of G. Note that the sums $s(A), s'(B)$ are all invariant under S and T. Thus *for any* G, A, A', B, B' we get the remarkable physical invariants

$$s(A)^*\,s(A'),\quad s(A)^* s'(B),\quad s'(B)^* s(A),\quad s'(B)^*\,s'(B'). \tag{4.7}$$

In the case of $\mathfrak{S}_3$ which we worked out in §3.3, $s(e) = s_2$, $s'(\psi_0) = s_5$, and $s'(\psi_0, \psi_1) = s_3$. This alone accounts for 9 of the 32 physical invariants for $\mathfrak{S}_3$.

Incidentally, taking A and B to be the trivial groups completes the proof of Proposition 1(h): the fifth physical invariant of course is the diagonal sum. For the second claim there, there is a natural degree-preserving bijection (see §3.5 of [3]) between the $\chi \in \mathrm{Irr}(G)$ with $\ker(\chi) \geq Z(G)$, and $\mathrm{Irr}(G/Z(G))$, so if $G/Z(G)$ is nonabelian there will always be degree > 1 characters appearing in each $s(A)$ and so they will be different from any $s'(B)$.

5 Twisting, Group Cohomology, and Projective Representations

As advertised in the introduction, one way to generalize the group data described in Section 2 is by introducing some "twisting". This twisting has a cohomological origin, as in the theory of affine algebras [29], where the infinitely many possible twists are labelled by the level k, an integer. In constrast, the twisting of the finite group modular data offers but a finite number of possibilities.

The twisting of this modular data was first described in the most generality by [4, 7, 20] (see also [6]), although to our knowledge explicit expressions for the modular matrices S and T in the most general case have not appeared until now (though the most general fusions appear in (6.44) of [4]). We will recall here their construction, as concretely as possible, and then we will compute explicit examples, making contact with known structures (most notably affine algebras).

5.1 Cohomological preliminaries

As a twisting of the finite group modular data is effected by elements of cohomology groups $H^i(G, U(1)) \cong H^i(G, \mathbb{C}^\times) \cong H^{i+1}(G, \mathbb{Z})$, we will first review the relevant properties and give some examples. For our purposes it is not very important to know how these are defined. We refer the reader to standard textbooks (like [38]) for a more complete treatment. For projective representations, see [2].

For all $i > 0$, $H^i(G, U(1))$ is a finite abelian group, which we will usually write additively, obeying $|G|H^i(G, U(1)) = 0$ (i.e. all elements have order dividing the order of G). For any G, $H^0(G, U(1)) = 0$, while $H^1(G, U(1)) = G/G'$ is the group of one-dimensional representations of G (G' is the commutator subgroup). The next group $M(G) := H^2(G, U(1))$, called the Schur multiplier (or multiplicator), classifies the (projectively) inequivalent projective representations, and for that reason, it will play a central role in the whole construction. The only other group we will be interested in is $H^3(G, U(1))$.

The following results can be useful: the square of the exponent of $M(G)$ divides $|G|$; if G is a p-group (i.e. its order is a power of a prime), then $H^3(G, U(1)) \neq 0$; if e is the exponent of G and e_i is the exponent of $H^i(G, U(1))$, then ee_2, e_1e_2 and e_2e_3 all divide $|G|$.

Computing cohomology groups, even $M(G)$, is usually difficult. They are known however for several familiar groups. For a cyclic group, one has $M(\mathbb{Z}_n) = 0$ and $H^3(\mathbb{Z}_n, U(1)) = \mathbb{Z}_n$. For products of identical cyclic groups,

$$M(\mathbb{Z}_n^k) = \mathbb{Z}_n^{k(k-1)/2}, \quad H^3(\mathbb{Z}_n^k, U(1)) = \mathbb{Z}_n^{k(k^2+5)/6}. \tag{5.1}$$

More generally, the Schur multiplier for any abelian group is as follows: write $G = \mathbb{Z}_{d_1} \times \mathbb{Z}_{d_2} \times \cdots \times \mathbb{Z}_{d_s}$ where $d_1|d_2|\cdots|d_s$, then

$$M(G) = \prod_{j=1}^{s} \mathbb{Z}_{d_j}^{j-1}. \tag{5.2}$$

Other known groups are (see e.g. Ch.6 of [3])

$$\begin{aligned}
&M(\mathfrak{S}_3) = 0, \quad H^3(\mathfrak{S}_3, U(1)) = \mathbb{Z}_6, \quad M(\mathfrak{S}_n) = \mathbb{Z}_2 \quad \text{for } n \geq 4, && (5.3)\\
&M(\mathfrak{A}_n) = \mathbb{Z}_2 \quad \text{for } n \geq 4,\ n \neq 6, 7, \quad M(\mathfrak{A}_n) = \mathbb{Z}_6 \quad \text{for } n = 6, 7, && (5.4)\\
&M(\mathfrak{D}_n) = 0, \quad H^3(\mathfrak{D}_n, U(1)) = \mathbb{Z}_{2n} \qquad \text{for } n \text{ odd}, && (5.5)\\
&M(\mathfrak{D}_n) = \mathbb{Z}_2, \quad H^3(\mathfrak{D}_n, U(1)) = \mathbb{Z}_n \times \mathbb{Z}_2 \times \mathbb{Z}_2 \qquad \text{for } n \text{ even}, && (5.6)\\
&M(\mathfrak{Q}_n) = 0, \quad H^3(\mathfrak{Q}_4, U(1)) = \mathbb{Z}_8, && (5.7)\\
&M(\mathrm{SL}_n(\mathbb{F}_q)) = 0, && (5.8)\\
&M(\mathrm{PSL}_n(\mathbb{F}_q)) = \mathbb{Z}_{\gcd(q-1,n)}, && (5.9)
\end{aligned}$$

where in (5.8) and (5.9), $(n, q) \notin \{(2, 4), (2, 5), (2, 7), (2, 9), (3, 2), (3, 3), (3, 4), (4, 2)\}$.

For odd prime p, there are precisely two different non-abelian groups of order p^3: one of these (a split extension of cyclic groups, the $p \neq 2$ analogue of $\mathfrak{D}_4$) has $M = 0$ and $H^3 = \mathbb{Z}_p \times \mathbb{Z}_p$; the other has $M = \mathbb{Z}_p \times \mathbb{Z}_p$ and $H^3 = \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$. (All groups of order p^2 are abelian.)

This gives some taste of what $M(G)$ and $H^3(G)$ look like for nice groups. The Schur multiplier is known for all simple groups. For instance it is trivial for the Monster.

As mentioned before, the Schur multiplier plays a central role in the theory of projective representations. Indeed a normalised 2-cocycle $\beta \in Z^2(G, U(1))$ is a map $G \times G \to U(1)$ satisfying $\beta(x, 1) = \beta(1, x) = 1$ and the cocycle condition $\beta(x, y)\,\beta(xy, z) = \beta(y, z)\,\beta(x, yz)$ for all $x, y, z \in G$. For any such β, one may consider the projective β-representations, i.e. the maps $\tilde{\rho} : G \to \mathrm{GL}(V)$ obeying $\tilde{\rho}(x)\tilde{\rho}(y) = \beta(x, y)\tilde{\rho}(xy)$. The cocycle condition corresponds to associativity. If β is identically 1, $\tilde{\rho}$ will be an ordinary representation and will be called *linear*.

If β, β' are cohomologous (i.e. $\beta^{-1}\beta'$ is a 2-coboundary – a *2-coboundary* is any $\beta \in Z^2$ of the form $\beta = \gamma(x)\gamma(y)\gamma(xy)^{-1}$), then any β'-representation will be projectively equivalent to some β-representation, and there will be the same number of k-dimensional β-representations as k-dimensional β'-representations, for each $k = 1, 2, \ldots$. Moreover, a β-representation can be one-dimensional only if β is a coboundary.

A natural question is to classify the projective representations belonging to a given cocycle. So let $r(G, \beta)$ denote the number of linearly inequivalent irreducible β-representations of G. We know that $r(G, \beta)$ is a cohomology class invariant. In order to compute this number, one introduces the notion of a β-regular group element: $g \in G$ is β-regular if $\beta(g, h) = \beta(h, g)$

for all h in $C_G(g)$. One may check that g is β-regular if and only if all its conjugates are, so that a whole conjugacy class is β-regular or not. Then

$$r(G,\beta) = \text{number of } \beta\text{−regular conjugacy classes of } G. \tag{5.10}$$

As an immediate consequence, one has the inequality $r(G,\beta) \leq k(G) = r(G,1)$ for any 2-cocycle β, with equality if (but not only if) β is a coboundary.

The projective characters $\tilde{\chi} \in \beta\text{-Irr}(G)$ (defined as usual by the trace of the representation) share many properties with the usual characters, except that they are in general not class functions. Schur's lemma still holds, from which orthogonality and completeness relations follow: for any $\tilde{\chi}, \tilde{\chi}' \in \beta\text{-Irr}(G)$ and any β-regular $a \in G$,

$$\frac{1}{|G|}\sum_{g\in G} \tilde{\chi}(g)^*\, \tilde{\chi}'(g) = \delta_{\tilde{\chi},\tilde{\chi}'}, \tag{5.11}$$

$$\frac{|K_a|}{|G|}\sum_{\tilde{\chi}\ \beta-\mathrm{Irr}} \tilde{\chi}(a)^*\, \tilde{\chi}(b) = \delta_{b\in K_a}\,. \tag{5.12}$$

The dimensions of the irreducible projective representations divide the order of the group, and their squares sum up to $|G|$. All β-characters vanish at non-regular classes.

Examples of projective representation data include the following. Writing $H^2(\mathbb{Z}_n^2, U(1)) = \langle\beta_1\rangle \cong \mathbb{Z}_n$ and $H^2(\mathbb{Z}_n^3, U(1)) = \langle\beta_1,\beta_2,\beta_3\rangle \cong \mathbb{Z}_n^3$, one finds

$$r(\mathbb{Z}_n^2, \beta_1^x) = [\mathrm{GCD}(x,n)]^2, \qquad r(\mathbb{Z}_n^3, \beta_1^x\,\beta_2^y\,\beta_3^z) = n\cdot[\mathrm{GCD}(x,y,z,n)]^2. \tag{5.13}$$

for any $1 \leq x,y,z \leq n$. For the dihedral groups, one has

$$r(\mathfrak{D}_{2n+1},\beta) = n+2, \qquad r(\mathfrak{D}_{2n},\beta) = \begin{cases} n+3 & \text{if } \beta \text{ is coboundary,} \\ n & \text{otherwise.}\end{cases} \tag{5.14}$$

These numbers are known also for $\mathfrak{S}_n$ and $\mathfrak{A}_n$.

5.2 General construction

The possible twistings are parametrised by the finitely many elements $\alpha \in H^3(G,U(1))$. The 3-cocycle condition is $\alpha(g,h,k)\,\alpha(g,hk,\ell)\,\alpha(h,k,\ell) = \alpha(gh,k,\ell)\,\alpha(g,h,k\ell)$. Different 3-cocycles give rise to different modular data, with in general different numbers of primary fields (never more than in the untwisted case), and with matrices S^α and T^α.

The construction starts from a normalised[§] element α of $H^3(G,U(1))$. We can and will assume for convenience that the values of α are always roots of 1 – in fact if α has cohomological order n, then we can require α to take the values of nth roots of 1 (proof: Write $\alpha^n = \delta\beta$; choose any $\gamma = \beta^{-1/n}$, then $\alpha\gamma$ is cohomologous to α and has the desired property.) Because $|G|\,H^3(G,U(1)) = 0$, n will necessarily divide $|G|$.

For all $a,g,h \in G$, we define auxiliary quantities

$$\beta_a(h,g) = \alpha(a,h,g)\,\alpha(h,h^{-1}ah,g)^{-1}\,\alpha(h,g,(hg)^{-1}ahg). \tag{5.15}$$

[§] That is, it satisfies $\alpha(e,h,g) = \alpha(h,e,g) = \alpha(h,g,e) = 1$. This implies that all 2-cocycles are accordingly normalised, $\beta_e(h,g) = \beta_h(e,g) = \beta_h(g,e) = 1$ for all h,g.

It follows from this definition that the β_a's are normalised twisted cocycles on G, namely they satisfy

$$\beta_a(x,y)\,\beta_a(xy,z) = \beta_a(x,yz)\,\beta_{x^{-1}ax}(y,z), \qquad \forall x,y,z \in G. \tag{5.16}$$

Furthermore, the restriction of each β_a to $C_G(a)$ clearly is a normalised 2-cocycle. As such they define projective representations of $C_G(a)$.

The primary fields Φ^α in the model twisted by a given 3-cocycle α will consist of all pairs $(a,\tilde{\chi})$ where $a \in R$ (as before) and $\tilde{\chi} \in \beta_a$-$\mathrm{Irr}(C_G(a))$. As a consequence of the inequality $r(H,\beta) \le k(H)$, we find

$$|\Phi^\alpha| = \sum_{a \in R} r(C_G(a), \beta_a) \le |\Phi|. \tag{5.17}$$

A simplification seems to occur when the 3-cocycle α is such that each 2-cocycle β_a is a coboundary on $C_G(a)$, which is the case considered in [6]. We call the resulting twisting "CT" (cohomologically trivial). Note that when $M(C_G(a)) = 0$ for all a, any twist α will automatically be CT. By no means though are CT twistings restricted to the groups whose centralisers have trivial Schur multipliers — see §6.2 for a class of examples. A CT-twisted theory has the same number of primaries as the untwisted one considered in the previous sections (but different modular matrices).

For CT twistings, [6, 4] proposed explicit formulae for the modular matrices that resemble the untwisted case except with additional phases. Each β_a being a coboundary on $C_G(a)$ by hypothesis, (5.16) can be used to show[¶] that one can find 1-cochains $\epsilon_a : C_G(a) \to U(1)$ for which $\epsilon_a(e) = 1$ and both

$$\beta_a(h,g) = (\delta\epsilon_a)(h,g) = \epsilon_a(h)\,\epsilon_a(g)\,\epsilon_a(hg)^{-1}, \tag{5.18}$$

$$\epsilon_{x^{-1}ax}(x^{-1}hx) = \frac{\beta_a(x, x^{-1}hx)}{\beta_a(h,x)}\epsilon_a(h), \tag{5.19}$$

for all $g,h \in C_G(a)$ and $x \in G$. Now, if $\tilde{\rho}$ is a β_a-representation with character $\tilde{\chi}$, then clearly $\rho(g) = \epsilon_a^{-1}(g)\tilde{\rho}(g)$ is a linear representation with character $\chi = \epsilon_a^{-1}\tilde{\chi}$. The modular matrices then become [6]

$$S^\alpha_{(a,\chi),(a',\chi')} = \frac{1}{|C_G(a)| \cdot |C_G(a')|} \sum_{g \in G(a,a')} \chi^*(ga'g^{-1})\,\chi'^*(g^{-1}ag)\,\sigma^*(a|ga'g^{-1}), \tag{5.20}$$

$$T^\alpha_{(a,\chi),(a',\chi')} = \delta_{a,a'}\delta_{\chi,\chi'}\frac{\chi(a)}{\chi(e)}\,\epsilon_a(a), \tag{5.21}$$

where the function $\sigma(\cdot|\cdot)$ is

$$\sigma(h|g) = \epsilon_h(g)\,\epsilon_g(h). \tag{5.22}$$

(Since $g \in C_G(h)$ if and only if $h \in C_G(g)$, this definition makes sense.) It is easy to see that (5.20),(5.21) are (essentially, i.e. up to a relabelling of the characters χ) independent of the choice of 1-cochains ϵ_a in (5.18),(5.19). These equations permit CT twists to be analysed as thoroughly as the untwisted data.

The interpretation involving 3-cocycles was developed in [4]. Though they give the fusion coefficients for arbitrary α, neither they nor to our knowledge anyone else has given S^α and

---

[¶] We thank Peter Bántay for correspondence on this point.

T^α for nonCT α, *explicitly* in terms of quantities directly associated with G (the closest is [11], which gives S^α in terms of $D^\alpha(G)$ characters). The formulae in the most general case can be obtained by going back to first principles.

The quantum double $D^\alpha(G)$ is a finite dimensional quasi-Hopf quasi triangular algebra. Its representations form a braided monoidal modular category with explicitly known universal braiding morphisms R_{12} , R_{21}. The point is that (up to normalisation) the mapping class group representations, in our case the matrices S and T, can be identified with Markov traces of intertwiners defined from coloured ribbon links. In particular, S_{ij} corresponds to the i, j coloured Hopf link and T_i to the twist θ (or v^{-1} in the [20] convention). We then obtain (as explained in the Appendix)

$$\begin{aligned} S^\alpha_{(a,\tilde\chi),(b,\tilde\chi')} &= S^\alpha_{(e,1),(e,1)} \,\mathrm{Tr}_{(a,\tilde\chi),(b,\tilde\chi')} \Big(R_{21}R_{12}\Big)^* \\ &= \frac{1}{|G|} \sum_{g\in K_a, g'\in K_b\cap C_G(g)} \Big(\frac{\beta_g(g',x^{-1})\beta_{g'}(g,y^{-1})}{\beta_g(x^{-1},h)\beta_{g'}(y^{-1},h')}\Big)^* \,\tilde\chi(h)^*\,\tilde\chi'(h')^*, \\ &= \frac{1}{|G|} \sum_{g\in K_a, g'\in K_b\cap C_G(g)} \Big(\frac{\beta_a(x,g')\beta_a(xg',x^{-1})\beta_b(y,g)\beta_b(yg,y^{-1})}{\beta_a(x,x^{-1})\beta_b(y,y^{-1})}\Big)^* \,\tilde\chi(h)^*\,\tilde\chi'(h')^*, \end{aligned} \tag{5.23}$$

where $g = x^{-1}ax = y^{-1}h'y,\ g' = y^{-1}by = x^{-1}hx,\ h \in C_G(a),\ h' \in C_G(b)$, and

$$T^\alpha_{(a,\tilde\chi),(b,\tilde\chi')} = \delta_{a,b}\delta_{\tilde\chi,\tilde\chi'}\frac{\tilde\chi(a)}{\tilde\chi(e)}\,T^\alpha_{(e,1),(e,1)}, \tag{5.24}$$

where $T^\alpha_{(e,1),(e,1)}$ can equal any third root of 1 (i.e. c is a multiple of 8), as for the untwisted data. The normalisation $S^\alpha_{(e,1),(e,1)} = \frac{1}{|G|}$ can be obtained from the orthogonality relations (5.11),(5.12) of projective characters.

The phases β_g here are defined by (5.15). Note that they are evaluated in (5.23) on elements which are not in $C_G(a)$ or $C_G(b)$. Unfortunately this makes the derivation of (5.20) from (5.23) more difficult — see question (1) in §7.

If we let n be the cohomological order of α, then the β_a will also have nth roots of 1 as its values. Thm. 6.5.15 in [2] then implies that the projective β_a-characters will have values in $\mathbb{Q}[\xi_{n\,e(G)}]$. Hence the entries of S^α and T^α will also lie in that field. Note that this field specialises to $\mathbb{Q}[\xi_{e(G)}]$ in the untwisted ($n = 1$) case, which was our previous result.

That S^α and T^α depend only on the cohomology class $\alpha \in H^3(G, U(1))$, is clear from the $D^\alpha(G)$ interpretation. A direct derivation of this for CT α is sketched in [4].

5.3 Analysis of twisted modular data

It is important to realise that the 2-cocycle β_e in (5.15) is identically 1. Thus $(e, \chi) \in \Phi^\alpha$ for any $\chi \in \mathrm{Irr}(G)$, and we obtain the useful formula

$$S^\alpha_{(e,\chi),(b,\tilde\chi')} = \frac{1}{|C_G(b)|}\tilde\chi'(e)\,\chi(b)^* \tag{5.25}$$

We immediately see from this that once again all quantum dimensions will be integers, and $(a, \tilde{\chi}) \in \Phi^\alpha(G)$ will be a simple current iff $a \in Z(G)$ and $\tilde{\chi}$ is degree-one (which implies β_a is coboundary).

Rationality of the entries $S_{(e,1),(b,\tilde{\chi}')}$ implies all Galois parities $\epsilon_\sigma(a, \tilde{\chi}) = +1$, and also that the vacuum $(e, 1)$ is fixed by all Galois automorphisms, exactly as before.

From (5.25) we also learn about the Galois action on arbitrary primaries. Choose $\ell \in \mathbb{Z}^\times_{|G|^2}$, then $\sigma_\ell(a, \tilde{\chi}) = (a_\ell, \tilde{\chi}_\ell)$, where a_ℓ denotes the element in $K_{a^\ell} \cap R$ as before, and where $\tilde{\chi}_\ell$ is some projective β_{a_ℓ}-character with dimension equal to that of $\tilde{\chi}$. To see this, consider

$$\chi(a_\ell)^* = \sigma_\ell \frac{S^\alpha_{(e,\chi),(a,\tilde{\chi})}}{S^\alpha_{(e,1),(a,\tilde{\chi})}} = \sigma_\ell \chi(a)^* = \chi(a^\ell)^* \tag{5.26}$$

for all $\chi \in \mathrm{Irr}(G)$, and hence $a_\ell \in K_{a^\ell}$. In particular, specialising to $\ell = -1$ tells us about charge conjugation.

Proposition 1(a),(c),(e),(f) are thus exactly as before. As mentioned before, the order of T^α will divide ne, where n is the order of α, so in particular the order of T^α will always divide $|G|^2$. We also see directly from (5.23) that $S^\alpha_{(a,\tilde{\chi}),(b,\tilde{\chi}')} = 0$ unless $K_b \cap C_G(a)$ has β_a-regular elements and $K_a \cap C_G(b)$ has β_b-regular elements. Thm. 1 becomes

Theorem 3. Let S and T be the Kac-Peterson matrices corresponding to an affine algebra $X_r^{(1)}$ at some level $k \geq 1$ (where X_r is simple). Let G be a finite group with $S^\alpha(G) = S$ and $T^\alpha(G) = \varphi T$ for some third root φ of 1. Then either:

(i) as before, either $(X_r, k) = (E_8, 1)$ and $G = \{e\}$, or $(X_r, k) = (D_{8n}, 1)$ and $G = \mathbb{Z}_2$;

(ii) $(X_r, k) = (A_{n^2-1}, 1)$ and $G = \mathbb{Z}_n$, n odd, for a specific twist;

(iii) $(X_r, k) = (B_{(m^2-1)/2}, 2)$ and $G = \mathfrak{D}_m$, m odd, for a specific twist.

The proof is very similar to that of Theorem 1, the only difference being the specific handling of the finitely many algebras and levels which survive the Galois and quantum dimension arguments. For example, use the formulae $S_{00} = \frac{1}{\sqrt{r+1}}$ and $c = r$ for $A_r^{(1)}$ level 1. The demonstration of *(ii)* and *(iii)* is made explicit in section 6.

We would also like an analogue of Thm.2. This is more difficult, but a key observation is that $|\Phi^\alpha| \geq 2k(G) - 1$. Indeed, $\beta_e \equiv 1$ gives us $k(G)$ primaries of the form (e, χ), and the remaining $k(G) - 1$ conjugacy classes K_a will each contribute $r(a, \beta_a) \geq 1$ primaries. Also important is the observation made in §6.2 that β_a will always be coboundary when $C_G(a) \cong \mathbb{Z}_n^2$ for some n.

Theorem 4. The only groups with at most 20 primaries are $\mathbb{Z}_1$, $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathfrak{S}_3$, $\mathfrak{A}_4$, $\mathbb{Z}_4$, $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathfrak{D}_5$, $\mathfrak{S}_4$, $\mathfrak{D}_4$, and the order 48 Frobenius group $\mathbb{Z}_2^4 \times_f \mathbb{Z}_3$ (defined in [35]), with at least 1, 4, 9, 8, 14, 16, 16, 16, 18, 19, and 19 primaries.

The proof is like that of Thm.2, using the tables of [35]. For instance, we know it is sufficient to consider up to class number $k(G) = 8$. Consider e.g. $G = \mathfrak{S}_4$: it has $k(G) = 5$ and centralisers $\mathfrak{S}_4$, $\mathfrak{D}_4$, $\mathbb{Z}_4$, $\mathbb{Z}_2 \times \mathbb{Z}_2$, and $\mathbb{Z}_3$. Thus a lower bound for $|\Phi^\alpha|$ is $5 + 2 + 4 + 4 + 3 = 18$.

In §3.3 we gave a basis in terms of which the untwisted S and T become permutation matrices. The analogue of this for the twisted data is as follows [11]. Let $C^\alpha(G_{comm})$ denote the space of all functions $f : G \times G \to \mathbb{C}$ for which $f(a, b) = 0$ unless a and b commute, and which obey the formula

$$f(x^{-1}ax, x^{-1}bx) = \frac{\beta_a(x, x^{-1}bx)}{\beta_a(b, x)} f(a, b) \tag{5.27}$$

Then S and T act on $C^\alpha(G_{comm})$ by $(Sf)(a, b) = \beta_b(a, a^{-1})^* f(b, a^{-1})$ and $(Tf)(a, b) = \beta_a(a, b)\, f(a, ab)$. See Questions 5 and 7 of §7. Incidentally, this action is completely natural in the case of twisted partition functions which involve projective representations of the symmetry group in some twisted sectors, and follows from the action of the modular group on the homological cycles of the torus — see [39] for examples in WZW models.

6 Twisted examples

We give in this final section a few examples of twisted finite group modular data, in order to give a flavour as to how they differ from the untwisted ones.

6.1 Abelian cyclic groups

One can easily illustrate the previous formalism in the case of a cyclic group $G = \mathbb{Z}_n$. Write $\langle\cdot\rangle : \mathbb{Z} \to \{0, 1, \ldots, n-1\}$ for reduction modulo n. One has $H^3(\mathbb{Z}_n, U(1)) = \mathbb{Z}_n$, and the following explicit representatives

$$\alpha_q(g_1, g_2, g_3) = \exp\Big\{2\mathrm{i}\pi q g_1(g_2 + g_3 - \langle g_2 + g_3\rangle)/n^2\Big\}, \tag{6.1}$$

where $q \in \mathbb{Z}_n$ parametrizes the different classes.

One easily computes that $\beta_a(h, g) = \alpha_q(a, h, g)$ is a coboundary for every q (expected since $M(\mathbb{Z}_n) = 0$). One finds $\sigma(h|g) = e^{4\mathrm{i}\pi qhg/n^2}$.

The linear characters of $\mathbb{Z}_n$ are $\chi_\ell(a) = e^{2\mathrm{i}\pi a\ell/n}$ for $\ell \in \mathbb{Z}_n$, and the modular matrices take the simple forms

$$S^{\alpha_q}_{(a,\chi_\ell),(a',\chi_{\ell'})} = \frac{1}{n}\,\exp\Big\{-2\mathrm{i}\pi[2qaa' + n(a\ell' + a'\ell)]/n^2\Big\}, \tag{6.2}$$

$$T^{\alpha_q}_{(a,\chi_\ell),(a,\chi_\ell)} = \exp\Big\{2\mathrm{i}\pi[qa^2 + na\ell]/n^2\Big\}. \tag{6.3}$$

Note that the charge conjugation $C = S^2$ is given $C(0, \chi_\ell) = (0, \chi_{-\ell})$ and $C(a, \chi_\ell) = (n - a, \chi_{-\ell-2q})$ if $a \neq 0$.

The identity corresponds to $(0, \chi_0)$, and as in the untwisted case, all primary fields are simple currents. The Verlinde formula yields the fusion coefficients

$$N^{(a'',\chi_{\ell''})}_{(a,\chi_\ell),(a',\chi_{\ell'})} = \delta_{a'',\langle a+a'\rangle}\,\delta_{\ell'',\ell+\ell'+2q(a+a'-a'')/n}. \tag{6.4}$$

This is in agreement with the charge conjugation, if one thinks of the conjugate of a field ϕ as the unique field ϕ^* such that the fusion $\phi \times \phi^*$ contains the identity. The fusion group (the group of simple currents) is isomorphic to $\mathbb{Z}_f \times \mathbb{Z}_{n^2/f}$ with $f = \mathrm{GCD}(2q, n)$.

The entries of S and T lie in $\mathbb{Q}(\xi_{n^2})$, which has Galois group $\mathbb{Z}_{n^2}^\times$. It is not difficult to see that the Galois action on the primaries is by fusion powers (as is always the case for simple currents):

$$\sigma_h(a,\chi_\ell) = (a,\ell)^{\times h}, \qquad h \in \mathbb{Z}_{n^2}^\times. \tag{6.5}$$

The physical invariants are known in all cases, since all primaries are simple currents. Their number varies much with the value of q (of f). Two extreme cases are $f=1$ ('maximal' twisting) for which the number of physical invariants is equal to $\sigma_0(n^2)$, the number of divisors of n^2 [40], and $f=n$ (no twisting), for which their number is $2(n+1)$ if n is odd prime [27].

We close this simple example by showing that it gives, for a specific twisting, the affine modular data of su(n^2), level 1, if n is an odd integer.

One may first make a few simple observations. The central charge of su$(n^2)_1$ is equal to $c \equiv n^2-1 \equiv 0 \pmod 8$ for n odd. The affine quantum dimensions $S_{0,j}/S_{0,0}=1$ all equal 1, so all affine primaries are simple currents, and form a group isomorphic to $\mathbb{Z}_{n^2}$. Thus one may hope for a relation with twisted $G=\mathbb{Z}_n$ data where the twist obeys $f=\mathrm{GCD}(q,n)=1$.

The affine primaries can be labelled by integers $j=0,1,2,\ldots,n^2-1$ modulo n^2. The affine matrices S and T are

$$S^{\mathrm{aff}}_{j,j'} = \frac{1}{n}\, e^{2\mathrm{i}\pi jj'/n^2}, \tag{6.6}$$

$$T^{\mathrm{aff}}_{j,j} = T_{0,0}\, e^{2\mathrm{i}\pi j(n^2-j)/2n^2}, \tag{6.7}$$

where $T_{0,0}=e^{2\mathrm{i}\pi c/24}$ is some third root of unity, which we will ignore.

It is now a simple matter to see that the two sets of matrices exactly coincide provided one chooses the twisting parameter as $q=\frac{n^2-1}{2}$. The bijection $j \leftrightarrow (a,\chi_\ell)$ between the two sets of primary fields is given by $j=a-n\ell$. Note that for this specific value of q, the fusion coefficients (6.4) amount to the addition modulo n^2.

6.2 Abelian non-cyclic groups

The simplest non-cyclic group is $G=\mathbb{Z}_n^2$, but it leads to nothing really new. The cohomology group $H^3(\mathbb{Z}_n^2,U(1))=\mathbb{Z}_n^3$ has three generators, but all 3-cocycles α lead to β's which are all coboundaries. Thus all twistings are CT, despite the fact that the Schur multiplier $M(\mathbb{Z}_n^2)=\mathbb{Z}_n$ is not trivial. (This fact was important for the proof of Thm.4.)

More interesting is the case $G=\mathbb{Z}_n^3$, for which $M(\mathbb{Z}_n^3)=\mathbb{Z}_n^3$ and $H^3(\mathbb{Z}_n^3,U(1))=\mathbb{Z}_n^7$. Following [41, 42], the generators of H^3 can be taken to be (same notations as above, the group elements are triplets $a=(a_1,a_2,a_3)$)

$$\alpha_I^{(j)}(a,b,c) = \exp\Big\{2\mathrm{i}\pi a_j(b_j+c_j-\langle b_j+c_j\rangle)/n^2\Big\}, \qquad 1\le j\le 3, \tag{6.8}$$

$$\alpha_{II}^{(jk)}(a,b,c) = \exp\Big\{2\mathrm{i}\pi a_j(b_k+c_k-\langle b_k+c_k\rangle)/n^2\Big\}, \qquad 1\le j<k\le 3, \tag{6.9}$$

$$\alpha_{III}(a,b,c) = \exp\Big\{2\mathrm{i}\pi a_1b_2c_3/n\Big\}. \tag{6.10}$$

An arbitrary 3-cocycle is a monomial in the generators, but only those which involve a non-trivial power of α_{III} define non-CT twistings. In other words, all α which contain a fixed cocycle of type III give rise to 2-cocycles β_a which are cohomologically equivalent, and hence lead to theories with the same number of primaries. In order to give a first feeling for non-CT twistings, we will compute the number of primary fields. It is sufficient to take α of type III, namely $\alpha = \alpha_{III}^q$, for $q \in \mathbb{Z}_n$. The 2-cocycles one obtains are then

$$\beta_a(b,c) = \exp\Big\{2\mathrm{i}\pi q(a_1b_2c_3 - b_1a_2c_3 + b_1c_2a_3)/n\Big\}. \tag{6.11}$$

Given a, we want to count the number of classes b (elements here) which are β_a-regular, i.e. which satisfy $\beta_a(b,c) = \beta_a(c,b)$ for all c. Taking successively $c = (1,0,0), (0,1,0)$ and $(0,0,1)$, the β_a-regular elements b are those which satisfy

$$a_2b_3 - a_3b_2 \equiv a_1b_3 - a_3b_1 \equiv a_1b_2 - a_2b_1 \equiv 0 \pmod f, \tag{6.12}$$

where $f = n/\mathrm{GCD}(q,n)$. The number of solutions $(b_1,b_2,b_3) \in \mathbb{Z}_n^3$ to this modular linear system is equal to $n^3 \cdot [\mathrm{GCD}(a_1,a_2,a_3,f)/f]^2$, which is the result announced in (5.13).

It remains to sum those numbers for all a to obtain the number of primaries. The result is an arithmetical function, best expressed in terms of the prime decomposition of $f = \prod_p p^{k_p}$:

$$|\Phi^\alpha| = \frac{n^6}{f^3} \prod_{\substack{p|f \\ p \text{ prime}}} \Big[(p^{k_p} - 1)(1 + p^{-1} + p^{-2}) + 1\Big]. \tag{6.13}$$

The modular matrices can be given quite explicitly in the general case, for all n and for any type of 3-cocycle. However, they are complicated arithmetic functions of the various parameters, something that obscures the structure. To simplify, we consider here the case when n is an odd prime number, and when the 3-cocycle is α_{III}.

When G is abelian, all factors in the formula (5.23) for S^α that involve the cocycles drop out, and we are left with the simple expressions:

$$S^\alpha_{(a,\tilde\chi),(b,\tilde\chi')} = \frac{1}{|G|}\tilde\chi^*(b)\,\tilde\chi'^*(a), \qquad T^\alpha_{(a,\tilde\chi),(b,\tilde\chi')} = \delta_{a,b}\,\delta_{\tilde\chi,\tilde\chi'}\,\frac{\tilde\chi(a)}{\tilde\chi(e)}, \tag{6.14}$$

where $\tilde\chi$ and $\tilde\chi'$ are respectively β_a- and β_b-projective characters, for the cocycles given above in (6.11) with $q = 1$.

It remains to compute the projective characters. To simplify, consider n an odd prime. One then finds n inequivalent irreducible β_a-projective representations of dimension n if a is not the identity, while there are of course n^3 representations of dimension 1 if $a = e$. Depending on the value of $a = (a_1,a_2,a_3)$, the characters are given in the following table, where it is implicit that the element $g = (g_1,g_2,g_3)$ must be β_a-regular for the character not to vanish. In the first three cases, the character label u runs over $\mathbb{Z}_n$, and in the last column, $\vec u$ takes all values in $\mathbb{Z}_n^3$.

	$a_1 \neq 0$	$a_1 = 0,\ a_2 \neq 0$	$a_1 = a_2 = 0,\ a_3 \neq 0$	$a_1 = a_2 = a_3 = 0$
$\tilde\chi(g)$	$n\,\xi_n^{a_1^{-1}ug_1 - a_1^{-1}a_2a_3g_1^2/2}$	$n\,\xi_n^{a_2^{-1}ug_2}$	$n\,\xi_n^{a_3^{-1}ug_3}$	$\xi_n^{\vec u\cdot\vec g}$

The condition that g must be β_a-regular makes the components of a and g play a symmetrical role. If a_2 is also invertible for instance, then (6.12) yields $a_1^{-1}g_1 = a_2^{-1}g_2$, so that the first character value is also equal to $\tilde{\chi}(g) = n\,\xi_n^{a_2^{-1}ug_2 - a_1a_2^{-1}a_3g_2^2/2}$.

The primary fields are thus $(e, \chi_{\vec{u}})$ and $(a, \tilde{\chi}_u)$, for a total of $|\Phi^\alpha| = n^3 + (n^3-1)n = n^4 + n^3 - n$.

The formulae for S and T are now straightforward to establish. Taking the condition of β-regularity into account, one finds that S is almost block-diagonal:

$$S^\alpha_{(a,\tilde{\chi}_u),(b,\tilde{\chi}_{u'})} =$$
$$\frac{1}{n}\begin{pmatrix} \frac{1}{n^2} & \frac{1}{n}\xi_n^{-u_1b_1-u_2b_2-u_3b_3} & \frac{1}{n}\xi_n^{-u_2b_2-u_3b_3} & \frac{1}{n}\xi_n^{-u_3b_3} \\ \frac{1}{n}\xi_n^{-u_1'a_1-u_2'a_2-u_3'a_3} & \begin{matrix}\xi_n^{-ua_1^{-1}b_1-u'b_1^{-1}a_1+(a_2a_3b_1+b_2b_3a_1)/2} \\ \times\delta(b_2-a_1^{-1}a_2b_1)\,\delta(b_3-a_1^{-1}a_3b_1)\end{matrix} & 0 & 0 \\ \frac{1}{n}\xi_n^{-u_2'a_2-u_3'a_3} & 0 & \begin{matrix}\xi_n^{-ua_2^{-1}b_2-u'b_2^{-1}a_2} \\ \times\delta(b_3-a_2^{-1}a_3b_2)\end{matrix} & 0 \\ \frac{1}{n}\xi_n^{-u_3'a_3} & 0 & 0 & \xi_n^{-ua_3^{-1}b_3-u'b_3^{-1}a_3} \end{pmatrix}, \tag{6.15}$$

where the blocks correspond to the subsets $\{a=e\}$, $\{a_1 \neq 0\}$, $\{a_1 = 0,\ a_2 \neq 0\}$, and $\{a_1 = a_2 = 0,\ a_3 \neq 0\}$.

The T matrix is particularly simple

$$T_{(a,\tilde{\chi}),(a,\tilde{\chi})} = \begin{cases} 1 & \text{for } a = e = (0,0,0), \\ \xi_n^{u-a_1a_2a_3/2} & \text{otherwise.} \end{cases} \tag{6.16}$$

One may check that they satisfy the expected relations $S^2 = (ST)^3 = C$ with the charge conjugation given by $C(e,\chi_{\vec{u}}) = (e, \chi_{-\vec{u}})$ and $C(a,\tilde{\chi}_u) = (a^{-1}, \tilde{\chi}_{u-a_1a_2a_3})$ for $a \neq e$. More generally, the Galois transformations on the primary fields take a somewhat unusual form[‖]

$$\sigma_\ell(a,\tilde{\chi}_u) = \begin{cases} (e, \chi_{\ell\vec{u}}) & \text{for } a = e, \\ (a^\ell, \tilde{\chi}_{\ell^2u+a_1a_2a_3\ell^2(\ell-1)/2}) & \text{for } a \neq e, \end{cases} \qquad \ell \in \mathbb{Z}_n^\times. \tag{6.17}$$

All Galois parities are equal to $+1$. One also checks the relation $\sigma_{\ell^2}T_{(a,\tilde{\chi}),(a,\tilde{\chi})} = T_{\sigma_\ell(a,\tilde{\chi}),\sigma_\ell(a,\tilde{\chi})}$.

Finally the fusion algebra can be computed, which shows a structure radically different from the untwisted or CT case. Here too, the fusion coefficients generically involve the group law and the structure constants for the irreducible characters. Since the latter have degree 1 or n, the fusion of two primary fields may contain 1, n or n^2 primary fields, with possible multiplicities (they turn out to be 1 or n only). The resulting formulae are tedious to write in full form as various cases need be distinguished. As an illustration, we give the coefficients

---

[‖] The unusual factor ℓ^2 in front of u (instead of ℓ) is a consequence of the insertion of factors a_i^{-1} in front of u in the projective characters. Those insertions are purely conventional and help make the symmetry in the components a_i manifest.

when $a, b, c \neq e$:

$$N^{(c,\tilde{\chi}_{u''})}_{(a,\tilde{\chi}_u),(b,\tilde{\chi}_{u'})} = \delta(a{+}b{-}c) \begin{cases} 1 \quad \text{for } c_1 = 0 \text{ and } a_1, b_1 \neq 0, \\ \Big[n\,\delta(a \text{ is } \beta_b\text{-reg})\,\delta(a_1^{-1}u + b_1^{-1}u' + a_2 b_3 - c_1^{-1}u'') \\ \qquad + \delta(a \text{ is not } \beta_b\text{-reg}) \Big] \quad \text{for } c_1 \neq 0 \text{ and } a_1, b_1 \text{ not both } 0, \\ \Big[n\,\delta(a \text{ is } \beta_b\text{-reg})\,\delta(a_2^{-1}u + b_2^{-1}u' - c_2^{-1}u'') \\ \qquad + \delta(a \text{ is not } \beta_b\text{-reg}) \Big] \quad \text{for } a_1, b_1 = 0 \text{ and } a_2, b_2 \neq 0, \\ \Big[n\,\delta(a \text{ is } \beta_b\text{-reg})\,\delta(a_3^{-1}u + b_3^{-1}u' - c_3^{-1}u'') \\ \qquad + \delta(a \text{ is not } \beta_b\text{-reg}) \Big] \quad \text{for } a_1, a_2, b_1, b_2 = 0 \text{ and } a_3, b_3 \neq 0. \end{cases} \tag{6.18}$$

An intriguing observation made in [42] (what he called *electric/magnetic duality*) is that the modular data for $\mathbb{Z}_2^3$ twisted by α_{III} *equals* that of untwisted $\mathfrak{D}_4$ (for an appropriate identification of primary fields), while the twist $\alpha_I^{(1)}\alpha_{III}$ yields the $\mathfrak{Q}_4$ modular data. We will address this again in Question **(8)** in §7.

6.3 Odd dihedral groups

The simplest non-abelian groups are the dihedral groups $\mathfrak{D}_m$. Their Schur multipliers are equal to 0 or $\mathbb{Z}_2$ for m odd or even respectively. For m odd, all centralisers of elements of $\mathfrak{D}_m$ also have trivial Schur multipliers, implying that all twistings of the modular data will be CT. Despite that fact, we will precisely consider m odd here[**], since we want to show that a particular CT twisting of the $\mathfrak{D}_m$ modular data yields nothing but the affine modular data of the odd orthogonal series B_ℓ at level 2, and for $m = \sqrt{2\ell+1}$.

The relevant group theoretic data for $\mathfrak{D}_m$ have been recalled in section 3.2. The number of primary fields is the same for all twistings, and given by $|\Phi| = \frac{m^2+7}{2}$.

The third cohomology group $H^3(\mathfrak{D}_m, U(1)) = \mathbb{Z}_{2m}$ is cyclic and so all 3-cocycles are powers of some generator. They have been very explicitly determined in [41]. Write the elements of $\mathfrak{D}_m$ as $g = s^A\,r^a$ for $A = 0, 1$ and $a = 0, 1, \ldots, m-1$. The group law then takes the form

$$(A, a)(B, b) = (\langle A + B\rangle_2, \langle (-1)^B a + b\rangle_m) \tag{6.19}$$

where the notation $\langle x\rangle_n$ means taking the residue of x modulo n between 0 and $n-1$. Then the 3-cocycles are given by [41]

$$\alpha((A, a), (B, b), (C, c)) = \exp\left\{ -\frac{2\mathrm{i}\pi p}{m^2} \Big[(-1)^{B+C} a[(-1)^C b + c - \langle (-1)^C b + c\rangle_m] + \frac{m^2}{2} ABC \Big] \right\}, \tag{6.20}$$

where $p = 0, 1, 2, \ldots, 2m-1$ labels the cohomology classes.

We now proceed to compute the matrices S and T from the formulas (5.20) and (5.21). We first compute

$$G(e, g) = \mathfrak{D}_m, \quad G(r^k, r^l) = \mathfrak{D}_m, \quad G(s, s) = \{e, s\}, \quad G(s, r^a) = \emptyset. \tag{6.21}$$

[**] The specific case $m = 3$ has been treated in [6].

In particular, the last equality implies $S^\alpha_{(s,\chi),(r^k,\chi')} = 0$. The remaining entries of S^α only require knowing the following values of σ, which easily follow from (5.15), (5.18) and (5.22),

$$\sigma(e|g) = 1, \qquad \sigma(s|s) = (-1)^p, \qquad \sigma(r^a|r^b) = \exp\left(\frac{-4\mathrm{i}\pi pab}{m^2}\right). \tag{6.22}$$

The rest is just a matter of a few calculations. We will write the two modular matrices with their rows and columns indexed by

$$(e,\psi_0), (e,\psi_1), (s,1), (s,\chi \neq 1), (e,\chi_i), (r^k,\chi_\gamma), \tag{6.23}$$

where χ_γ are characters of the group $\mathbb{Z}_m$, with values $\chi_\gamma(a) = \mathrm{e}^{-2\mathrm{i}\pi a\gamma/m}$. The indices i and γ run from 1 to $\frac{m-1}{2}$ and m respectively.

One finds

$$S^\alpha = \frac{1}{2m}\begin{pmatrix} 1 & 1 & m & m & 2 & 2 \\ 1 & 1 & -m & -m & 2 & 2 \\ m & -m & (-1)^p m & (-1)^{p+1} m & 0 & 0 \\ m & -m & (-1)^{p+1} m & (-1)^p m & 0 & 0 \\ 2 & 2 & 0 & 0 & 4 & 4\cos\frac{2\pi i\gamma}{m} \\ 2 & 2 & 0 & 0 & 4\cos\frac{2\pi i\gamma}{m} & 4\cos\frac{2\pi(2pkk'+\gamma k'+\gamma' k)}{m^2} \end{pmatrix}, \tag{6.24}$$

and

$$T^\alpha = \mathrm{diag}\Big[1, 1, \mathrm{e}^{4\mathrm{i}\pi p/8}, -\mathrm{e}^{4\mathrm{i}\pi p/8}, 1, \mathrm{e}^{2\mathrm{i}\pi(pk^2-mk\gamma)/m^2}\Big]. \tag{6.25}$$

On the affine side, the alcôve of B_ℓ level 2 contains $\ell+4$ primary fields, corresponding to the weights: 0, $2\omega^1$, $\omega^1+\omega^\ell$, ω^ℓ and ν^j, defined as $\nu^j = \omega^j$ for $1 \leq j \leq \ell-1$, and $\nu^\ell = 2\omega^\ell$. The conformal central charge is equal to 2ℓ, which is multiple of 8 if ℓ is a multiple of 4. For simplicity, we will use the height variable, $n = 2\ell+1$.

The affine S matrix, with the primary fields labelled as above, is equal to [43]

$$S = \frac{1}{2\sqrt{n}}\begin{pmatrix} 1 & 1 & \sqrt{n} & \sqrt{n} & 2 \\ 1 & 1 & -\sqrt{n} & -\sqrt{n} & 2 \\ \sqrt{n} & -\sqrt{n} & \sqrt{n} & -\sqrt{n} & 0 \\ \sqrt{n} & -\sqrt{n} & -\sqrt{n} & \sqrt{n} & 0 \\ 2 & 2 & 0 & 0 & 4\cos\frac{2\pi jk}{n} \end{pmatrix}, \tag{6.26}$$

while the affine T is

$$T = T_{0,0} \cdot \mathrm{diag}\Big[1, 1, -\mathrm{e}^{2\mathrm{i}\pi\ell/8}, \mathrm{e}^{2\mathrm{i}\pi\ell/8}, \mathrm{e}^{\mathrm{i}\pi j(n-j)/n}\Big]. \tag{6.27}$$

The comparison of the two pictures is now easy. As far as the first four primary fields are concerned, the two sets of modular data coincide provided $m = \sqrt{n} = \sqrt{2\ell+1}$ and $p \equiv 0 \pmod 2$ (the individual identifications among the third and fourth fields depend on the value of $\ell \pmod 8$). It implies that the finite group and the affine theories have the same number of primaries, $\ell + 4 = \frac{m^2+7}{2}$.

The remaining ℓ fields, on the affine side, are the ν^j. They can be set in correspondence with the finite group primary fields (e,χ_i) and (r^k,χ_γ), depending on whether j is a multiple of m or not, via the following formula:

$$\nu^{im}\sim(e,\chi_i),\qquad \nu^{k+m\gamma}\sim(r^k,\chi_\gamma). \tag{6.28}$$

When i,k and γ run over their domain, the index j indeed takes all integer values from 1 to $\ell=\frac{n-1}{2}$.

One may then check that the two sets of modular matrices coincide for a specific choice of p. For $j=i\sqrt{n}$, one has $j(n-j)\equiv 0 \pmod{2n}$, while for $j=k+\sqrt{n}\gamma$, one finds

$$j(n-j)\equiv\tfrac{n-1}{2}k^2-k\gamma\sqrt{n}\equiv pk^2-mk\gamma \pmod{2n}, \tag{6.29}$$

provided one makes the choice $p\equiv\frac{n-1}{2} \pmod{2n}$. It gives the unique twisting for which the finite group modular data and the affine data coincide. That the S matrices are equal is done in the same way.

Hence infinitely many physical invariant classifications for CT-twisted nonabelian G were done in [33]. If we let $d=\sigma_0(m^2)-1$ denote the number of divisors $d'\le m$ of m^2, then the number of physical invariants of $\mathfrak{D}_m$ for the given twist α is $d(d+3)/2+4$. Since $\mathfrak{D}_3\cong\mathfrak{S}_3$, we get 'only' 9 physical invariants for twisted $\mathfrak{S}_3$ – still a large number considering the small number of primaries, but dwarfed by the 32 ones for the untwisted data.

7 Questions

We conclude by collecting a small number of the questions raised in this paper.

(1)[††] We have been unable to derive the CT formula (5.20), appearing in e.g. [6], from our general formula (5.23), except in special cases such as when G is abelian. Though this isn't strictly necessary, it would make a nice consistency check of the modular data. Our formula (5.23) for S^α has the strength that it is given explicitly in terms of quantities directly associated with G, but it has the weakness that it involves the 2-cocycles β_a away from $C_G(a)$, and this is what complicates our attempt to derive (5.20). The derivations in [4, 11] involve considering the β_a's at the centralisers only.

(2) It should be possible to generalise the observations in §4.2 by considering the relations of the modular data for G with that of normal subgroups and quotient groups. This should yield an analogue of "conformal embeddings" for this data.

(3) Can we recover the character table from the matrices S and T? Do non-isomorphic groups have different S and T? These are big questions. Probably the answer to the first is yes, and to the second is no. It turns out to be very difficult to identify a group from its character table together with nice additional information – perhaps the modular data provides a means? A natural place to look for nonisomorphic groups with identical modular data are the order 16 groups with identical group algebras, or Brauer pairs (i.e. groups with identical character tables and identical power maps $K_a\mapsto K_{a^n}$).[‡‡] Perhaps relevant is [44], which proves that two groups have the same character table iff their group algebras are isomorphic as quasi-Hopf algebras.

[††] We thank Peter Bántay for correspondence regarding this point.

[‡‡] We thank John McKay for bringing Brauer pairs to our attention.

(4) We see in places a tantalising hint of some sort of duality between the group element component of the primary field, and its character component (see e.g. §4.2). Is there any way to make this precise?

(5) Does the $\mathrm{SL}_2(\mathbb{Z})$ representation obtained from the finite group modular data, factor through a congruence subgroup? Presumably the answer is always yes. Certainly it is true for untwisted data. The easiest way to see this uses the permutation-like representation of [6, 11, 12], so a natural starting point for the twisted data would involve the basis given at the end of §5.3.

(6) An important fact for affine modular data is the Kac-Peterson formula, which provides a useful interpretation for the eigenvalues $S_{\lambda\mu}/S_{0\mu}$ of the fusion matrices. It would be highly desirable to find out what form if any that takes here.

(7) What is the analogue for the affine modular data of the basis discussed at the end of §5.3, in which S and T appear as generalised permutation matrices?

(8) As mentioned at the end of §6.2, twisted $\mathbb{Z}_2^3$ modular data yields [42] that of untwisted $\mathfrak{D}_4$ and $\mathfrak{Q}_4$ modular data. It is natural to expect that there are many more such examples: e.g. circumstantial evidence (such as quantum dimensions and numbers of primaries) suggests that appropriately twisted $\mathbb{Z}_p^3$ would yield the modular data for the Heisenberg group (consisting of all 3×3 upper triangular matrices with 1's down the diagonal and entries in $\mathbb{Z}_p$), or the order p^3 generalisation of $\mathfrak{D}_4$ (presentation $\langle a, b \,|\, a^{p^2} = b^p = b^{-1}aba^{-1-p} = e\rangle$). This is probably the analogue of "rank-level duality" for this finite group data. Note that it is not a coincidence that $\mathbb{Z}_2^3$ and $\mathfrak{D}_4$ both have the same order — if the matrices S are to be equal, their entries $S_{0,0}$ must agree and hence the two groups must have equal order. Again, perhaps relevant is [44], which may supply a starting point for a general theory.

Acknowledgements. T.G. thanks F. A. Bais, P. Bántay, J. McKay, D. Stanley, M. Walton, and B. Westbury; his research was supported in part by NSERC. A. C. thanks D. Altschuler, J. Lascoux, V. Pasquier, J.B. Zuber. This paper was initiated at Oberwolfach and partially written at IHES, and we thank both for their hospitality.

A Appendix

We sketch here the derivation of formula (5.24) for T^α. A basis of $D^\alpha(G)$ is labelled by pairs (g, x) of group elements with algebra product

$$(g, x) \cdot (k, y) \ := \delta_{g,xkx^{-1}} \ \beta_g(x, y) \ (g, xy). \tag{A.1}$$

Now, $D^\alpha(G)$ is semi-simple and finite-dimensional over $\mathbb{C}$, so its irreducibles $R(a, \tilde{\rho})$ are subrepresentations of the regular one. They can be described as follows [7]. The representation space for $R_{(a,\tilde{\rho})}$ is spanned by vectors $|x_j\rangle \otimes |w\rangle$, and the action is give by

$$R_{(a,\tilde{\rho})}(g, x)|x_j\rangle \otimes |w\rangle = \delta_{g,x_k a x_k^{-1}} \ \frac{\beta_g(x, x_j)}{\beta_g(x_k, h)} \ |x_j\rangle \otimes \ \tilde{\rho}(h)|w\rangle, \tag{A.2}$$

where x_j are coset representatives for $G/C_G(a)$ and $(x_k, h) \in G/C_G(a) \times C_G(a)$ are uniquely defined by $x\ x_j = x_k h$. Here $\tilde{\rho}$ denotes a projective irreducible β_a-representation of $C_G(a)$ obtained by restriction of the regular representation of $D^\alpha(G)$.

Contact with the regular representation has been made by Altschuler and Coste [20]. Let us give here an account of it. The central element $v^{-1} = \sum_{k \in G}(k, k)$ is easily diagonalised: For each conjuguacy class K_a of G denote by $e(a)$ the common order of its elements, then

$$\omega_a := \prod_{j=0}^{e(a)-1} \alpha(g, g^j x, x^{-1} g x) \tag{A.3}$$

is independent of the choice of $g \in K_a$ and $x \in G$.

Then all eigenvalues of v^{-1} are the $e(a)$-th roots of the ω_a's. For each such root λ, each $g \in K_a$ and each class $\{g^j x\}$ of $G/\langle g \rangle$, an eigenvector $\psi_\lambda[g, x]$ is easily constructed.

That these eigenvalues of projective β_a-representations are given in terms of (roots of) values of a 3-cocycle raises the interesting group theoretic question: is this a clever way of classifying or putting together all projective representations of a finite group in terms of 3-cocycles for extensions inside which it is a centraliser ?

The central element v is a constant on each $R_{(a,\tilde{\rho})}$, for convenience let us rather focus on v^{-1}, which has a simpler expression and on its eigenvalues: they are necessarily $e(a)$-th roots of the important quantities ω_a, numbers which should be considered as tabulation data of any 3-cocycle. All of their $e(a)$-th roots are eigenvalues of v^{-1}.

The following nice identity satisfied by the twisted non abelian 2-cocycle and valid for any $a, x \in G$, is due to Altschuler:

$$\beta_{xax^{-1}}(xax^{-1},\ x) = \beta_{xax^{-1}}(x,\ a) = \alpha(xax^{-1}, x, a) \tag{A.4}$$

and is proved by direct evaluation in terms of the 3-cocycle α. From it we get the remarkably simple expression: $R_{(a,\tilde{\rho})}(v^{-1})\ \ |x_j\rangle \otimes |w\rangle = \ |x_j\rangle \otimes\ \tilde{\rho}(a)|w\rangle = \lambda\ |x_j\rangle \otimes |w\rangle$ (here $|w\rangle$ spans the representation space of $\tilde{\rho}$. Because of the expression of v^{-1} the quotient of cocycles which appear in $R_{(a,\tilde{\rho})}$ becomes here independent of $x = x_j = x_k$ and in fact equal to 1.

Since for $h \in C_G(a)$,

$$\beta_a(h,\ a) = \beta_a(a,\ h) = \alpha(a, h, a) \tag{A.5}$$

we easily get, using an induction on j for $h = a^j$: $\tilde{\alpha}(a)^{e(a)} = \omega_a I$, $\tilde{\alpha}(a)$ central and equal to λI.

References

[1] G. Karpilovsky, *Group representations*, Vol. 1B, North-Holland, Amsterdam, 1992.

[2] G. Karpilovsky, *Projective representations of finite groups*, Marcel Dekker, New York, 1985.

[3] Y. G. Berkovich and E. M. Zhmud, *Characters of finite groups*, Part 1,2, AMS, Providence, 1998.

[4] R. Dijkgraaf and E. Witten, *Commun. Math. Phys.* **129**, 393 (1990).

[5] D. S. Freed and F. Quinn, *Commun. Math. Phys.* **156**, 435 (1993).

[6] R. Dijkgraaf, C. Vafa, E. Verlinde and H. Verlinde, *Commun. Math. Phys.* **123**, 485 (1989).

[7] R. Dijkgraaf, V. Pasquier and P. Roche, *Nucl. Phys. (Proc. Suppl.)* **18B**, 60 (1990).

[8] V.G. Kac and I. Todorov, *Commun. Math. Phys.* **190**, 57 (1997).

[9] P. Bántay, *Int. J. Mod. Phys.* **A9**, 1443 (1994).

[10] G. Mason, pp.405-417, in: *Groups '93*, Cambridge University Press, Cambridge, 1995.

[11] P. Bántay, *Lett. Math. Phys.* **22**, 187 (1991).

[12] T. Koornwinder, B. Schroers, N. Slingerland and F. A. Bais, J. Phys. A: Math. Gen. **32**, 8539 (1999).

[13] C. Dong and G. Mason, *Commun. Math. Phys.* **163**, 523 (1994).

[14] P. Bántay, *Phys. Lett.* **B419**, 175 (1998).

[15] F. A. Bais, P. van Driel, and M. de Wild Propitus, *Nucl. Phys.* **B393**, 547 (1993).

[16] G. Lusztig, pp.235-262, in: *Proc.Symp.Pure Math* **47**, AMS, Providence, 1987.

[17] G. Lusztig, *Duke Math. J.* **73**, 227 (1994).

[18] A. Yu. Kitaev, *Fault-tolerant quantum computation by anyons*, (preprint, quant-ph/9707021).

[19] C. Vafa, *Nucl. Phys.* **B273**, 592 (1986).

[20] D. Altschuler and A. Coste, *Commun. Math. Phys.* **150**, 83 (1992);
ibid., *J. Geom. Phys.* **11**, 191 (1993);
ibid., *R.C.P. 25, Vol. 46*, Prépubl. Inst. Rech. Math. Av., 1994/29, Univ. Louis Pasteur, Strasbourg, 1994 (and ref's therein).

[21] G. Kuperberg, *J. Knot Theory Ramifications* **5**, 173 (1996).

[22] Y. Kosaki, H. Munemasa and K. Yamagami, *Pacif. J. Math.* **177**, 269 (1997).

[23] P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory*, Springer, New York, 1997.

[24] A.N. Schellekens and S. Yankielowicz, *Nucl. Phys.* **B327**, 673 (1989).

[25] A. Coste and T. Gannon, *Phys. Lett.* **B323** (1994) 316-321.

[26] D. Bernard, *Nucl. Phys.* **B288**, 628 (1987).

[27] M. Kreuzer and A.N. Schellekens, *Nucl. Phys.* **B411**, 97 (1994).

[28] J. Fuchs, B. Gato-Rivera, A.N. Schellekens and C. Schweigert, *Phys. Lett.* **B334**, 113 (1994).

[29] V. Kac, *Infinite dimensional Lie algebras*, Cambridge University Press, Cambridge, 1990.

[30] M.A. Walton, *Phys. Lett.* **B241**, 365 (1990).

[31] J. Fuchs, *Commun. Math. Phys.* **136**, 345 (1991).

[32] A. Cappelli, C. Itzykson and J.-B.Zuber, *Commun. Math. Phys.* **113**, 1 (1987).

[33] T. Gannon, *The level 2 and 3 modular invariants for the orthogonal algebras*, (preprint, math.QA/9809020), to appear in *Canad. J. Math.*

[34] D.F. Holt and W. Plesken, *Perfect Groups*, Clarendon Press, Oxford, 1989.

[35] A. Vera López and J. Vera López, *Israel J. Math.* **51**, 305 (1985); *ibid* **56**, 188 (1986).

[36] T. Gannon, *Nucl. Phys.* **B491**, 659 (1997).

[37] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, 2nd edition, Springer, New York, 1992.

[38] K.S. Brown, *Cohomology of groups*, Springer, New York, 1982.

[39] S. Liénart, Ph. Ruelle and O. Verhoeven, in preparation.

[40] P. Degiovanni, *Commun. Math. Phys.* **127**, 71 (1990).

[41] M. de Wild Propitius, *Topological interactions in broken gauge theories*, Ph.D. thesis, University of Amsterdam, 1995.

[42] M. de Wild Propitius, *Nucl. Phys.* **B489**, 297 (1997).

[43] V.G. Kac and M. Wakimoto, *Adv. Math.* **70**, 156 (1988).

[44] A.A. Davydov, *Finite groups with the same character tables, Drinfel'd algebras and Galois algebras*, (preprint, math.RT/0001119).

Sketch of a Program for Universal Automorphic Functions to Capture Monstrous Moonshine

Igor FRENKEL and Robert PENNER

Institut des Hautes Études Scientifiques
35, route de Chartres
91440 – Bures-sur-Yvette (France)

Janvier 2021

IHES/M/21/01

SKETCH OF A PROGRAM FOR UNIVERSAL AUTOMORPHIC FUNCTIONS TO CAPTURE MONSTROUS MOONSHINE

IGOR FRENKEL AND ROBERT PENNER

Abstract. We review and reformulate old and prove new results about the triad $\mathrm{PPSL}_2(\mathbb{Z}) \subseteq \mathrm{PPSL}_2(\mathbb{R}) \circlearrowright ppsl_2(\mathbb{R})$, which provides a universal generalization of the classical automorphic triad $\mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathrm{PSL}_2(\mathbb{R}) \circlearrowright psl_2(\mathbb{R})$. The leading P or p in the universal setting stands for *piecewise*, and the group $\mathrm{PPSL}_2(\mathbb{Z})$ plays at once the role of universal modular group, universal mapping class group, Thompson group T and Ptolemy group. In particular, we construct and study new framed holographic coordinates on the universal Teichmüller space and its symmetry group $\mathrm{PPSL}_2(\mathbb{R})$, the group of piecewise $\mathrm{PSL}_2(\mathbb{R})$ homeomorphisms of the circle with finitely many pieces, which is dense in the group of orientation-preserving homeomorphisms of the circle. We produce a new basis of its Lie algebra $ppsl_2(\mathbb{R})$ and compute the structure constants of the Lie bracket in this basis. We define a central extension of $ppsl_2(\mathbb{R})$ and compare it with the Weil-Petersson form. Finally, we construct a $\mathrm{PPSL}_2(\mathbb{Z})$-invariant 1-form on the universal Teichmüller space formally as the Maurer-Cartan form of $ppsl_2(\mathbb{R})$, which suggests the full program for developing the theory of automorphic functions for the universal triad which is analogous, as much as possible, to the classical triad. In the last section we discuss the representation theory of the Lie algebra $ppsl_2(\mathbb{R})$ and then pursue the universal analogy for the invariant 1-form $E_2(z)dz$, which gives rise to the spin 1 representation of $psl_2(\mathbb{R})$ extended by the trivial representation. We conjecture that the corresponding automorphic representation of $ppsl_2(\mathbb{R})$ yields the bosonic CFT_2. Relaxing the automorphic condition from $\mathrm{PSL}_2(\mathbb{Z})$ to its commutant allows the increase of the space of 1-forms six-fold additively in the classical case and twelve-fold multiplicatively in our universal case. This leads to our ultimate conjecture that we can realize the Monster CFT_2 via the automorphic representation for the universal triad. This conjecture is also bolstered by the links of both the universal Teichmüller and the Monster CFT_2 theories to the three-dimensional quantum gravity.

Keywords: universal Teichmüller space, Thompson group T, loop algebra for sl_2, Weil-Petersson form, Maurer-Cartan form, Monstrous Moonshine, three dimensional quantum gravity

---

Date: December 31, 2020.

Introduction

The idea of studying a universal Teichmüller space that contains the union of images of all classical Teichmüller spaces goes back to Lipman Bers [1], who considered the group of all quasisymmetric mappings of the circle $\mathbb{S}^1$ to itself modulo the Möbius group $\mathrm{PSL}_2(\mathbb{R})$.

Further development of the classical Teichmüller spaces [23] and their applications to string theory led the second-named author to a new model [24] of the universal Teichmüller space based on the group, denoted $\mathrm{Homeo}_+(\mathbb{S}^1)$, of all orientation-preserving self-homeomorphisms of $\mathbb{S}^1$ with the compact-open topology modulo the Möbius group. The problem of providing a parametrization for the new larger model of universal Teichmüller space and the corresponding group of homeomorphisms of $\mathbb{S}^1$ was also resolved in [24] by identification of the latter group with the space $\mathcal{T}ess'$ of all ideal tesselations of the hyperbolic plane together with a choice of distinguished oriented edge, or simply doe, that is,

$$\mathrm{Homeo}_+(\mathbb{S}^1) \approx \mathcal{T}ess'. \tag{0.1}$$

In particular, the identity element in $\mathrm{Homeo}_+(\mathbb{S}^1)$ corresponds to a special tesselation called the Farey tesselation. It is obtained by applying the modular group $\mathrm{PSL}_2(\mathbb{Z})$ to the doe running from 0 to ∞. In view of the isomorphism (0.1), one can consider coordinates on $\mathcal{T}ess = \mathcal{T}ess'/\mathrm{PSL}_2(\mathbb{R})$ as a measure of the distortion of a given tesselation from the specified Farey tesselation by so called shearing coordinates associated to every nonoriented edge of the Farey tesselation, or equivalently to the elements of $\mathrm{PSL}_2(\mathbb{Z})/(\mathbb{Z}/2)$, where the $\mathbb{Z}/2$ subgroup reverses the orientation of edges. The shearing coordinates can be realized as logarithms of cross-ratios of certain hyperbolic lengths and are invariant under the action of the Möbius group. As a result we obtain an injection

$$\mathrm{Homeo}_+(\mathbb{S}^1)/\mathrm{PSL}_2(\mathbb{R}) \to \prod_{e\in\{\text{edges}\}} \mathbb{R}_+^e, \tag{0.2}$$

where $\mathbb{R}_+^e$ is a copy of $\mathbb{R}_+$ associated to the edge e.

Two questions arise in relation to the parametrization (0.2): the first is how to characterize its image, and the second is how to circumvent the factorization by the Möbius group and obtain directly the coordinates of $\mathrm{Homeo}_+(\mathbb{S}^1)$ itself. To answer these questions, we introduce new

---

It is a pleasure to acknowledge the hospitality of the Institut des Hautes Études Scientifiques, where this work began, and in particular the excellent assistance of Fanny Dufour with the figures.

holographic coordinates and framings in Sections 2 and 3 the paper. These new coordinates are actually based upon elaborations of the shearing coordinates defined on a decorated bundle from [23] over the universal Teichmüller space which was central also in [24].

The shearing coordinates as well as our holographic coordinates admit especially simple transformation under the dense subgroup $\mathrm{PPSL}_2(\mathbb{R}) \subseteq \mathrm{Homeo}_+(\mathbb{S}^1)$. This subgroup in its turn contains a discrete subgroup $\mathrm{PPSL}_2(\mathbb{Z}) \subseteq \mathrm{PPSL}_2(\mathbb{R})$ of piecewise $\mathrm{PPSL}_2(\mathbb{Z})$ homeomorphisms with rational breakpoints between pieces. Elements of $\mathrm{PPSL}_2(\mathbb{Z})$ turn out automatically to be once-continuously differentiable on the circle. This pair of groups contains the classical pair $\mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathrm{PSL}_2(\mathbb{R})$, which is the first hint towards the new extended theory of automorphic forms.

Another discrete group that is indispensable in the Teichmüller theory is the mapping class group or Teichmüller modular group. In our universal context it is realized by the group of the Farey-type tesselations which coincide with the Farey tesselation outside of a finite polygon and known as the Ptolemy group Pt, also introduced in [24].

More precisely, the *flip* on an edge e in a tesselation is defined by replacing it by the other diagonal of the quadrilateral complementary to $\cup(\tau - \{e\})$ in $\mathbb{D}$; if e is the doe, then the flip on it is enhanced by inducing the orientation coming from the counter-clockwise rotation of e. The *Ptolemy group(oid)* Pt has objects given by tesselations with doe of $\mathbb{D}$ which coincide with the Farey tesselation outside of a finite polygon and morphisms given by finite compositions of flips. Triangulations with doe are combinatorially rigid, and this allows flips to be labeled by edges of a fixed tesselation, so words in these labels render Pt in fact a group. Furthermore, $\mathrm{PSL}_2(\mathbb{Z})$ sits inside Pt as those tesselations which are identical to the Farey tesselation except perhaps for the location of the doe.

The remarkable fact of the universal setting is that under the isomorphism (0.1), these two discrete subgroups of $\mathrm{Homeo}_+(\mathbb{S}^1)$ coincide $\mathrm{PPSL}_2(\mathbb{Z}) \approx \mathrm{Pt}$. Furthermore, this universal mapping class group is also isomorphic to the celebrated Thompson group T, so

$$\text{Thompson T} \approx \mathrm{PPSL}_2(\mathbb{Z}) \approx \text{Ptolemy Pt}. \tag{0.3}$$

The rich combinatorial structure of T is studied in [5] and numerous sequels. In particular, T admits a presentation by means of two generators and certain relations similar to those of the modular group.

The classical theory of automorphic forms on $\mathrm{PSL}_2(\mathbb{R})$ involves, besides the classical modular group $\mathrm{PSL}_2(\mathbb{Z})$, also the large class of discrete subgroups $\Gamma \subseteq \mathrm{PSL}_2(\mathbb{R})$ commensurable with the modular group. In our universal setting, one considers a similar class of infinite discrete groups $\mathrm{P}\Gamma$ associated to such subgroups $\Gamma \subseteq \mathrm{PSL}_2(\mathbb{R})$. By definition, $\mathrm{P}\Gamma$ is the subgroup of $\mathrm{Homeo}_+(\mathbb{S}^1)$ consisting of piecewise Γ homeomorphisms with finitely many rational breakpoints. To study these groups, one can consider again the paving, i.e., decomposition into finite-sided congruent ideal polygons, determined by the action of Γ on the doe from 0 to ∞. In Section 1, we consider the special case when $\Gamma = \mathrm{PSL}_2(\mathbb{Z})'$ is the commutant of the modular group. We also review the results of [19] about the class of finitely generated such groups $\mathrm{P}\Gamma$, which turn out to be precisely the groups of genus zero. This is a hint towards Monstrous Moonshine, and we conclude Section 1 with the problem of characterizing those $\Gamma \subseteq \mathrm{PSL}_2(\mathbb{Z})$ that occur in Monstrous Moonshine in these terms.

In the classical theory of automorphic functions, besides the pair of groups $\mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathrm{PSL}_2(\mathbb{R})$, the Lie algebra $psl_2 = psl_2(\mathbb{R})$ plays a pivotal role; we employ this notation psl_2 in the current discussion instead of the more standard notation $sl_2 = sl_2(\mathbb{R})$ simply to emphasize the relationship with the associated Lie group. To develop the universal counterpart of the classical theory, one requires a suitable Lie algebra for the topological group $\mathrm{Homeo}_+(\mathbb{S}^1)$. It has been argued in [20] that this infinite-dimensional counterpart is precisely the algebra of piecewise sl_2 vector fields on the circle with finitely many pieces and rational breakpoints between them. This Lie algebra was denoted psl_2, where the where the p stood for *piecewise*. In the present paper, this Lie algebra will be denoted $ppsl_2 = ppsl_2(\mathbb{R})$ to emphasize its relationship with the group $\mathrm{PPSL}_2(\mathbb{R})$. In Sections 4 and 5, we continue the study of this Lie algebra $ppsl_2$. In particular following [20], we find a basis parametrized by the edges of the Farey tesselation, or equivalently by $\mathrm{PSL}_2(\mathbb{Z})/(\mathbb{Z}/2)$, and explicitly derive the commutation relations in this basis. In the next Section 6, we define the central extension of the Lie algebra $ppsl_2$ viewed as a loop algebra and compare it with the central extension given by the universal Weil-Petersson 2-form first studied in [24], which naturally extends the classical Weil-Petersson Kähler form computed in [25].

All three structures of our universal triad

$$\mathrm{PPSL}_2(\mathbb{Z}) \subseteq \mathrm{PPSL}_2(\mathbb{R}) \circlearrowright ppsl_2(\mathbb{R}) \tag{0.4}$$

are combined in our construction of an automorphic 1-form on the universal Teichmüller space in Section 7. This construction is one of the main results of the paper, and it is the first step towards our program of developing the theory of automorphic forms for the triad (0.4). Since $\mathrm{PPSL}_2(\mathbb{R})$ allows for irrational breakpoints, the triad (0.4) should more properly be written $\mathrm{PPSL}_2(\mathbb{R}) \supseteq \mathrm{PPSL}_2(\mathbb{Z}) \circlearrowright ppsl_2(\mathbb{R})$. One might simply restrict to rational breakpoints for $\mathrm{PPSL}_2(\mathbb{R})$ as well in order to ensure (0.4), but let us not dwell on this perhaps interesting detail.

Though the proposed new theory is expected to be of a higher level of complexity than the usual theory for the classical triad

$$\mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathrm{PSL}_2(\mathbb{R}) \circlearrowright psl_2(\mathbb{R}),$$

one can pursue the analogy with the classical case whenever possible. In particular, we argue that our automorphic 1-form on the universal Teichmüller space has its classical counterpart in the 1-form $E_2(z)dz$ on the hyperbolic plane, the universal cover of the classical modular curve, where $E_2(z)$ is the non-holomorphic weight two covariant Eisenstein series.

In Appendix B, we study its lift to an automorphic function on $\mathrm{PSL}_2(\mathbb{R})$ and show that the resulting lift generates an indecomposable representation of $psl_2(\mathbb{R})$ with a one-dimesional sub-representation. This indecomposable representation together with its conjugate, and common one-dimensional representation, can be characterized as the harmonic subspace of automorphic functions, i.e., it is annihilated by the Laplace operator. Though the realization of the holomorphic and antiholomorphic discrete series of weights $4, 6, 8, \dots$ by the automorphic functions is well known, the case of weight 2 has not appeared in the literature to the best of our knowledge. Yet it is exactly an analogue of this special automorphic form which arises in our construction of a $\mathrm{PPSL}_2(\mathbb{Z})$-invariant 1-form on universal Teichmüller space in the formal guise of the Maurer-Cartan form of $\mathrm{PPSL}_2(\mathbb{R})$.

Our example of this universal automorphic form opens a new program of research in this subject for the triad (0.4), as discussed in Section 8. One of the first challenges is to find an analogue of the indecomposable representation generated by the lift of the 1-form $E_2(z)dz$ in our universal setting. Our conjecture explained in Section 8 is that this is the infinite symmetric power $S^\infty V$ of the bosonic space that appears in the classical theory, so that $S^\infty V$ is the free bosonic field and is one of the simplest CFT_2.

The theory of automorphic forms instantly enlarges various spaces by relaxing the the automorphic property to smaller discrete groups. Thus, the passage from $\mathrm{PSL}_2(\mathbb{Z})$ to its commutant $\mathrm{PSL}_2(\mathbb{Z})'$ of index six increases the space V additively 6-fold. Similarly, we expect that replacing $\mathrm{PPSL}_2(\mathbb{Z})$ by $\mathrm{PPSL}_2(\mathbb{Z})'$ will yield a correspondingly multiplicative increase of $S^\infty V$. Thus, we expect that this mechanism will allow the construction of the Monstrous CFT_2 of [12], and in other words capture the Monster as in the title of this paper.

In fact at the end of Section 8, we explain this from the other end of the theory by recalling that Monstrous Moonshine allows one to attach a genus zero subgroup $\Gamma \subseteq \mathrm{PSL}_2(\mathbb{R})$ to every conjugacy class of the Monster group. Then one can construct as above a Γ-invariant 1-form $E_2^\Gamma(z)dz$, where $E_2^\Gamma(z)$ is again the non-holomorphic weight two covariant Eisenstein series. This can be lifted to a representation V^Γ of $\mathrm{PSL}_2(\mathbb{R})$ isomorphic to V but realized by a different space of automorphic functions. Then the universal analogue is conjectured to yield a twisted Monster representation Λ^Γ as in in [10], and the comparison of Λ and Λ^Γ should reveal the Monster in the universal automorphic realization. It will be interesting to see the correspondence between the Thompson-like groups $\mathrm{P}\Gamma \subseteq \mathrm{PPSL}_2(\mathbb{R})$ introduced in Section 1 and the Monster Moonshine groups $\Gamma \subseteq \mathrm{PSL}_2(\mathbb{R})$, both of which are genus zero and satisfy certain additional properties [6, 7]

Our conjecture on the relationship between automorphic forms on the universal Teichmüller space and the Monster CFT_2 is strongly supported by the links of both subjects to yet another: three-dimensional quantum gravity. In fact, the link with universal Teichmüller theory has a long history in the physics literature going back to [33]. More recently, a rigorous definition of the universal phase space of AdS_3 gravity was given in [29], and it was proven there that it can be identified with the cotangent bundle over the universal Teichmüller space in its Bers formulation. Thus, one could expect that the space of states of AdS_3 quantum gravity can be realized in some class of functions on the universal Teichmüller space. However, it was also understood in the physics literature that the Bers version of universal Teichmüller space must be enlarged to account for all states related to black holes. Indeed, such enlargement was an original motivation for [24].

It was argued in [35] that the general philosophy of $\mathrm{AdS}_3/\mathrm{CFT}_2$ correspondence suggests that the space of states of the simplest pure quantum gravity is precisely the Monster CFT_2 constructed in [12]. This correspondence was further supported by the explanation of Monstrous Moonshine, and in particular by the mysterious genus zero property, from the point of view of twisted state sums of three-dimensional quantum gravity [10]. However in spite of all these tantalizing observations, a rigorous mathematical theory of three-dimensional quantum gravity is still missing. Thus, we believe that the development of the theory of universal automorphic representations can help to build a Monstertrap for all the Moonshine and at the same time help us to understand the true nature of three-dimesional quantum gravity.

1. Farey tesselation and modular group

Let $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ denote the integers, rational, real and complex numbers, respectively, with $\hat{\mathbb{Z}} \subseteq \hat{\mathbb{Q}} \subseteq \hat{\mathbb{R}} \subseteq \hat{\mathbb{C}}$ denoting their respective one-point compactifications by the point ∞ at infinity. Set $i = \sqrt{-1}$.

The open unit disk $\mathbb{D}$ in the complex plane $\mathbb{C}$ is identified with the Poincaré disk model of the hyperbolic plane in the standard way, where the boundary unit circle $\mathbb{S}^1$ is identified with the circle at infinity. Also consider the upper half-plane $\mathcal{U} = \{z = x + iy \in \mathbb{C} : y > 0\}$. The *Cayley transform*

$$\begin{aligned} C : (\mathcal{U}, \hat{\mathbb{R}}) &\to (\mathbb{D}, \mathbb{S}^1) \\ s &\mapsto \frac{s-i}{s+i} \end{aligned}$$

induces an isomorphism of pairs.

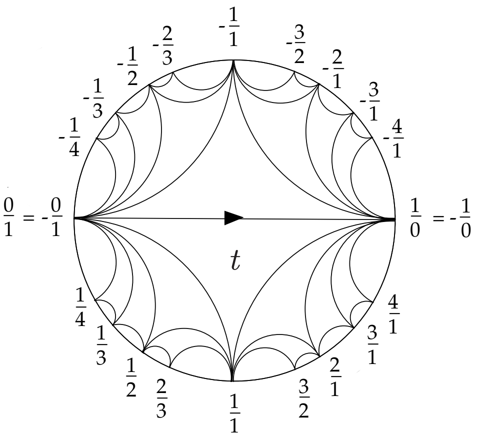


Figure 1. The first several generations of the Farey tesselation τ_* of the Poincaré disk $\mathbb{D}$ with its distinguished oriented edge.

Let t denote the ideal hyperbolic triangle with vertices $\pm 1, -i \in \mathbb{S}^1$ as in Figure 1, and consider the group $\mathcal{R}$ generated by reflections in the sides of t. Define the *Farey tesselation* τ_* to be the full $\mathcal{R}$-orbit of the frontier of t. τ_* has a distinguished oriented edge given by the interval from -1 to $+1$. A direct Euclidean construction of τ_* with a discussion of its history and number theoretic significance is given in [26, §3.1].

More generally, an arbitrary *tesselation* of $\mathbb{D}$ is a locally finite collection τ of hyperbolic geodesics decomposing $\mathbb{D}$ into complementary ideal triangles. Geodesics in τ are called its *edges*, and τ itself is regarded as a set of edges. Let $\tilde{\tau}$ denote the set of all oriented edges of τ, and if $e \in \tilde{\tau}$, then let $|e| \in \tau$ denote the unoriented edge underlying e. A *distinguished oriented edge* or *doe* on τ is the specification of an element $e \in \tilde{\tau}$. Let $\tau^0 \subseteq \mathbb{S}^1$ denote the set of all *vertices* of τ, namely, the collection of all endpoints of all edges in τ.

As is well-known, $\hat{\mathbb{Q}} \approx \tau_*^0$ under the Cayley transform, namely,

$$C\ (p/q) = \frac{p - iq}{p + iq} = \frac{p^2 - q^2}{p^2 + q^2} - i\ \frac{2pq}{p^2 + q^2} \in \mathbb{S}^1,$$

as indicated in Figure 1, and we shall refer to these points as the *rational points* of $\mathbb{S}^1$. We abuse notation slightly and sometimes let $\frac{p}{q} \in \hat{\mathbb{Q}}$ denote the point $C(p/q) \in \mathbb{S}^1$. $\frac{p}{q} \in \hat{\mathbb{Q}}$ is said to be of *generation* g if the radial arc in $\mathbb{D}$ from the origin to $\frac{p}{q}$ meets the interior of $g \geqslant 0$ distinct ideal triangles complementary to τ_*. The standard doe of τ_* runs between the two rational points of generation zero, from $\frac{0}{1}$ to $\frac{1}{0}$.

Another canonical tesselation of $\mathbb{D}$ with doe is the *dyadic tesselation* τ_d, which has the same doe as the Farey tesselation, and indeed the same generation one vertices as well, and which is recursively characterized by the property that one vertex of each triangle complementary to τ_d bisects the angle between its other two vertices. Thus, one has $\tau_d^0 = \{e^{2\pi i(1-k)} : k \in \mathbb{Z}\}$, the points in the circle with dyadic rational arguments. In contrast, τ_*^0 consists of points in the circle with rational rectilinear coordinates.

The *modular group*

$$\mathrm{PSL}_2 = \mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathcal{R}$$

of integral fractional linear transformations is the subgroup of $\mathcal{R}$ consisting of compositions of an even number of reflections, or in other words, the group of two-by-two integral matrices A of unit determinant modulo the equivalence relation generated by identifying A with $-A$. More generally, the *Möbius group*

$$M\ddot{o}b = \mathrm{PSL}_2(\mathbb{R}) \supseteq \mathrm{PSL}_2$$

consists of the two-by-two unimodular matrices over $\mathbb{R}$ modulo the same equivalence relation. *Möb* is the group of orientation-preserving hyperbolic isometries of $\mathbb{D}$.

In particular, $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2$ *acts on the right* (here following Gauss) on the rational points by

$$A : \frac{p}{q} \;\mapsto\; \frac{pd - qb}{qa - pc},$$

so the edge $e_A = (doe).A$ has initial point $-\frac{b}{a} = \frac{b+ia}{b-ia}$ and terminal point $-\frac{d}{c} = \frac{d+ic}{d-ic}$.

The *Thompson group* T is the collection of all orientation-preserving piecewise homeomorphisms of $\mathbb{S}^1$ with finitely many breakpoints among τ_d^0 which are affine in the coordinate θ on each piece. Recall from the Introduction the group $\mathrm{PPSL}_2(\mathbb{Z})$ of all piecewise $\mathrm{PSL}_2(\mathbb{Z})$ homeomorphisms of $\mathbb{S}^1$ with finitely many breakpoints among the Farey rationals τ_*^0. In fact, the Thompson group T is conjugate in $\mathrm{Homeo}(\mathbb{S}^1)_+$ to $\mathrm{PPSL}_2(\mathbb{Z})$, where the conjugating homeomorphism fixes the endpoints of the doe and maps the Farey tesselation to the dyadic tesselation of $\mathbb{D}$ in the natural way; this conjugating homeomorphism was first studied by Minkowski in [21] for its remarkable analytic properties.

Here are two standard propositions which are the starting points of our discussion:

Proposition 1.1. *The modular group* PSL_2 *leaves setwise invariant the Farey tesselation* τ_**, mapping* $\cup\tau_*$ *onto* $\cup\tau_*$*. Any orientation-preserving homeomorphism of the circle leaving invariant* τ_* *in this manner lies in* PSL_2*. The modular group acts simply transitively on* $\tilde{\tau}_*$*. A fundamental domain for the action of the modular group on* $\mathcal{U}$ *is given by* $\{x + iy \in \mathcal{U} : x^2 + y^2 > 1 \text{ and } |x| < \frac{1}{2}\}$. □

Proposition 1.2. *A generating set for* PSL_2 *is given by any pair of*

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix},\ S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

and $S^2 = 1 = R^3$ *is a complete set of relations in the generators* $R = T^{-1}U$ *and* $S = TU^{-1}T$*, so* $\mathrm{PSL}_2 \approx \mathbb{Z}/2 * \mathbb{Z}/3$*. In fact,* $T^{-1} = R^2S$ *and* $U = SR^2$ *are each of infinite order and are conjugate in* PSL_2. □

Complete relations in the generators U and T are given by $T^{-1}UT^{-1} = TU^{-1}T = UT^{-1}U = U^{-1}TU^{-1}$, with these so-called braiding relations reflecting the fact that PSL_2 is also the mapping class group of the once-punctured torus with U and T representing Dehn twists.

Geometrically, the elliptic element S setwise fixes $|e_I|$ and reverses the orientation of the doe e_I, R is the elliptic transformation cyclically permuting the vertices of the triangle to the right of the doe, and U (respectively T) is the parabolic transformation with the fixed point $\frac{0}{1}$ (respectively $\frac{1}{0}$) which cyclically permutes the incident edges of τ_* in the counter-clockwise sense about $\frac{0}{1}$ (respectively the clockwise sense about $\frac{1}{0}$). Typical aspects of our enumeration of oriented edges by elements of PSL_2 are illustrated in Figure 2.

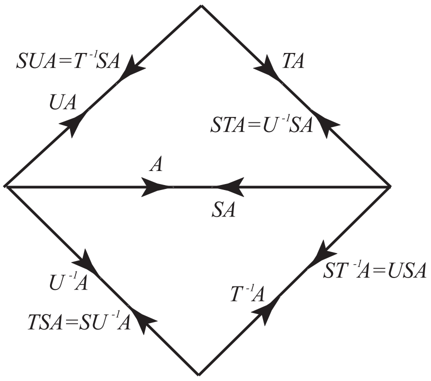


Figure 2. Enumeration of oriented edges near $|e_A| \in \tau_*$.

We turn our attention now to the *commutant* PSL_2' of PSL_2, namely, the subgroup of PSL_2 generated by the group commutators $aba^{-1}b^{-1}$, for $a, b \in \mathrm{PSL}_2$. As follows immediately from Proposition 1.2, the quotient $\mathrm{PSL}_2/\mathrm{PSL}_2' \approx \mathbb{Z}/6$.

Proposition 1.3. *Consider an element of* PSL_2 *written as a word* $W = W(U,T)$ *in the generators* U *and* T. *Then the coset of the element* W *in* $\mathrm{PSL}_2/\mathrm{PSL}_2'$ *is given by the residue modulo six of the total exponent sum of* W *in the letters* U *and* T^{-1}.

Proof. Each of U and T^{-1} is infinite order and maps to the generator of $\mathbb{Z}/6 \approx \mathrm{PSL}_2/\mathrm{PSL}_2'$. Since U and T^{-1} are conjugate in PSL_2, the word W may be written as a product of conjugates of U, which thus abelianizes to the exponent sum of U in this expression of W times the generator of $\mathbb{Z}/6$. □

Corollary 1.4. *Fix any triangle complementary to* $\cup\tau_*$ *and consider the six possible orientations on its frontier edges. The labels in* PSL_2 *of these six oriented edges span the six commutant cosets. Furthermore,* PSL_2' *corresponds to the collection* $\mathcal{C} \subseteq \tilde{\tau}_*$ *of oriented edges determined by the following conditions: the doe lies in* $\mathcal{C}$*; suppose an oriented edge*

$e \in \tilde{\tau}_*$ lying in $\mathcal{C}$ has ideal endpoint $v \in \mathbb{S}^1$, then every third edge of τ_* incident on v also lies in $\mathcal{C}$ with alternating orientations around v.

Proof. For the first part, consideration of Figure 2 shows that for the triangle to the right of the doe, we must show that the PSL_2'-cosets of $I, S, US, T^{-1} = T^{-1}, TS, U^{-1}$ are distinct. These have respective expressions in U, T given by I, $UT^{-1}U$, $U^2T^{-1}U$, T^{-1}, $TUT^{-1}U$, U^{-1} with respective total exponent sums in U and T^{-1} given by $0, 3, 4, 1, 2$, $-1 \equiv 5$. The general case follows upon conjugation.

For the second part, we likewise argue for the doe with the general result then following by conjugation. For the doe we must prove that $SU^{\pm 1}, SU^{\pm 2} \notin \mathcal{C}$ and $SU^{\pm 3} \in \mathcal{C}$ (since obviously $U^{\pm 1}, U^{\pm 2} \notin \mathcal{C}$) and likewise for T instead of U. To this end, we write $SU = UT^{-1}U^2$, $SU^{-1} = UT^{-1}$, $SU^2 = UT^{-1}U^3$, $SU^{-2} = UT^{-1}U^{-1}$, $SU^3 = UT^{-1}U^4$ and $SU^{-3} = UT^{-1}U^{-2}$ with respective total exponent sums in U and T^{-1} given by $4, 2, 5, -1, 6 \equiv 0, 0$ as required. □

The theory of the Ptolemy-Thompson group can be compared in richness only with the modular group. It is natural to ask what is the wider class of similar groups that have similar properties as $\mathrm{PPSL}_2(\mathbb{Z})$. In the case of the modular group, such a class is embodied into so-called arithmetic groups, namely, subgroups of $\mathrm{PSL}_2(\mathbb{R})$ which are commensurable with $\mathrm{PSL}_2(\mathbb{Z})$. They have an important common property with the modular group: their ideal compactification points in $\mathbb{D}/\Gamma$ are again $\hat{\mathbb{Q}}$. Thus, the piecewise Γ subgroups of $\mathrm{PPSL}_2(\mathbb{R})$, denoted by $\mathrm{P}\Gamma$, for Γ arithmetic, might serve as a natural generalization of $\mathrm{PPSL}_2(\mathbb{Z})$.

The first question one might ask about these groups $\mathrm{P}\Gamma$ is when they are finitely generated. The answer is contained in the dissertation of Laget [19]: if and only if Γ is of genus zero. But this is a hint towards a possible connection with Monstrous Moonshine, where the genus zero property also plays a key role. The groups Γ that appear in Monstrous Moonshine also have the width 1 property, namely the Farey-type tesselation or paving obtained by the action of Γ on the edge $(0, \infty)$ in $\mathbb{D}$ is 1-periodic.

These three conditions on Γ of arithmeticity, genus zero property and width 1 restrict the number of possible Γ to the large but finite number 6486 according to [8]. On the other hand, there are 194 conjugacy classes in the Monster, which give rise to groups Γ with the same three properties as above, with some identifications that reduce the number of groups Γ involved in Monster Moonshine to 171 [6]. The problem that we want to address in this conclusion to the section is:

How to characterize the Monster Moonshine Γ among arithmetic, genus 0, width 1 groups from the properties of the piecewise Γ groups $\mathrm{P}\Gamma$.

So the Monster began to show up already in Thompson-like groups. But to capture it, we shall need other aspects of our universal automorphic triad: The universal Teichmüller space and the corresponding Lie algebra which we shall review and study in the next sections.

2. Framed holographic coordinates

We shall regard $ppsl_2$ as an appropriate limit of copies of sl_2, one copy at each Farey rational point in $\mathbb{S}^1$. To take the limit, we require suitable coordinates on ideal polygons in which to calculate representations, and this section is dedicated to this end. Let us begin with the basic "lambda length" coordinates (sometimes called Penner coordinates) which we recall from [23, 26].

A *decorated ideal n-gon* is an ideal polygon of n sides in hyperbolic space together with n horocycles, one centered at each of its ideal vertices. One such coordinate is associated to each of the $2n-3$ edges in an ideal triangulation of the n-gon including its frontier edges. As in the Introduction, there is a basic move, called a *flip*, on the interior edge of such a triangulation, where one removes the edge so as to produce a complimentary ideal quadrilateral and then replaces it with the other diagonal of this quadrilateral to finally produce another ideal triangulation.

Theorem 2.1. *Fix some $n \geqslant 3$ and consider a* decorated *ideal n-gon P in $\mathbb{D}$. Suppose that the frontier edges of P are labeled and choose an ideal triangulation Δ of P. Then the moduli space of such decorated polygons up to the natural action of Möb is given by the assignment of one real lambda length $\lambda = \sqrt{\exp \ell}$ to each unoriented edge $e \in \Delta$, where ℓ is the signed hyperbolic distance along e between the horocycles centered at its endpoints, taken with a positive sign if and only if these horocycles are disjoint. Moreover, the lambda lengths are governed by the Ptolemy equation*

$$ef = ac + bd$$

whenever f arises from a flip on e in the quadrilateral bounded by a, b, c, d in this counter-clockwise cyclic order, here identifying an edge with its lambda length for convenience. □

Beyond the lambda lengths, other basic coordinates are the *h-lengths* assigned to any ideal vertex v of a decorated polygon P by taking the hyperbolic distance along the horocycle centered at v between the incident frontier edges of P. A fundamental formula illustrated in Figure 3 relates lambda lengths and h-lengths in a triangle, where $\frac{\lambda_i}{\lambda_j\lambda_k}$ is the h-length opposite the edge λ_i, for $\{i,j,k\}=\{0,1,2\}$, and we here and hereafter again often conflate an edge with its lambda length for convenience. It follows that the product of h-lengths of consecutive vertices is the reciprocal square of the lambda length of the edge they span, i.e., $\big(\frac{\lambda_i}{\lambda_j\lambda_k}\big)\big(\frac{\lambda_j}{\lambda_i\lambda_k}\big)=\frac{1}{\lambda_k^2}$, so either the triple of lambda lengths λ_i, for $i=0,1,2$ or the triple of h-lengths $\frac{\lambda_i}{\lambda_j\lambda_k}$, for $\{i,j,k\}=\{0,1,2\}$, give coordinates on the moduli space of $M\ddot{o}b$-orbits of decorated ideal triangles by the previous result.

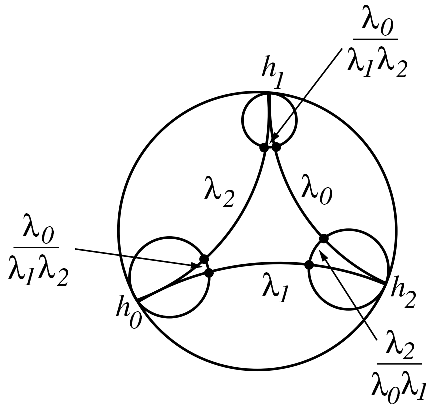

$\frac{\lambda_0}{\lambda_1\lambda_2}$
h_1
λ_2
λ_0
$\frac{\lambda_0}{\lambda_1\lambda_2}$
λ_1
h_0
h_2
$\frac{\lambda_2}{\lambda_0\lambda_1}$

Figure 3. Three horocycles h_i whose centers span a decorated ideal triangle with opposite lambda lengths λ_i and adjacent h-lengths $\frac{\lambda_i}{\lambda_j\lambda_k}$ for $\{i,j,k\}=\{0,1,2\}$.

Corollary 2.2. [Holographic parameters on a decorated polygon]: *For any $n\geqslant 3$, the moduli space of decorated ideal n-gons P up to the action of $M\ddot{o}b$ is parametrized by the assignment of the lambda length of each frontier edge together with the assignment of the h-length of each vertex of P. These parameters are constrained by three rational equations.*

Proof. The proof is by induction on n, and the basis step $n=3$ was just discussed. For the induction, choose consecutive edges a,b of P with intermediate vertex w, let T be the triangular convex hull of $\{a,b\}$, α,β the respective h-lengths in T opposite the endpoints of a,b distinct from w, and let γ denote the h-length at w of T. Consider also the h-lengths α',β' of the same vertices but in P rather than T.

Now consider the closure Q of the complement of T in P. Thus Q is an ideal $(n-1)$-gon which inherits all the moduli of P *except* that the respective h-lengths in Q at u and v are taken to be $\alpha'-\alpha$ and $\beta'-\beta$. The inductive hypothesis therefore applies to Q with these parameters, and hence also to P with parameters α', β' at u, v. Notice the constraint that $\alpha\beta$ equals the reciprocal square of the lambda length of $T \cap Q$. This constraint in T accounts for dimension reduction to confirm that the three rational constraints in Q persist in P. □

The significance of this parametrization arises upon consideration of the *Ptolemy groupoid* $\mathrm{Pt}(P)$ of an ideal polygon P whose objects are the ideal triangulations of P with interior doe, and whose morphisms are the flips along interior edges, where the flip on the doe e induces an orientation on the resulting edge and is of order four, whereas the flip on any other edge is of order two; these are the *face relations.* The flips on any two edges which do not lie in the frontier of a common triangle commute, and these are the *commutativity relations.* The *pentagon relations* arise from the serial sequence of flips alternating between two edges lying in the frontier of a common triangle, which has order five, unless one of these edges is the doe, in which case the order is ten. See [23, 24, 26] for further details.

According to Proposition 4.5 in [24] and remarks in [23, 24], we have

Theorem 2.3. *Finite sequences of flips supported on the interior of an ideal polygon P act transitively on the collection of all its triangulations with interior doe. It follows that flips generate* $\mathrm{Pt}(P)$. *Furthermore, a complete set of relations are given by the face, commutativity and pentagon relations.* □

So here finally is the significance of the holographic parameters: Taking the quotient by flips evidently renders meaningless the notion of lambda lengths of interior edges, these edges being precisely the data are obliterated. However, the lambda lengths of its frontier edges and the h-lengths of its vertices survive flips on interior edges to give parameters on the quotient. This explains our mechanism of holography.

We must go a bit further still in order to derive useful coordinates. An ordered triple of pairwise distinct points in the circle is a member of the configuration space

$$C_3 = \{(u, v, w) \in (\mathbb{S}^1)^3 : u, v, w \text{ are pairwise distinct}\},$$

on which *Möb* acts simply transitively according to Proposition 1.1. A *framing on an ideal polygon* is the specification of an arbitrary element of C_3, whose members need not be among the vertices of the polygon. A framed ideal polygon does not necessarily contain any doe, though an interior doe in a triangulation evidently determines an *associated framing* given by the respective initial and terminal points of the doe followed by the third vertex of the triangle immediately to the right of the doe.

An ordered triple $(u, v, w) \in C_3$ uniquely determines a unit tangent vector in $\mathbb{D}$ as follows. There is a unique geodesic g asymptotic to u and v, and the orthogonal projection of w onto g provides a point $p \in \mathbb{D}$. The unit tangent vector to g at p, where g is oriented with w on its right, finally provides the asserted unit vector associated to (u, v, w). In these coordinates on the unit tangent bundle to $\mathbb{D}$, the almost complex structure is furthermore conveniently described by $(u, v, w) \mapsto (w, x, v)$, where $\mathbb{S}^1 \ni x \neq w$ is the endpoint of the geodesic through p which is asymptotic to w.

Remark 2.4. The term “framing” is motivated by instantons, where the incorporation of framing dramatically simplifies the formulas and accounts for the gauge group, cf. [3]. The same phenomena occur here: the framing kills both the $M\ddot{o}b$ action and the three relations on holographic parameters in order to produce finally useful coordinates.

To explain the utility of the new coordinates, let us consider lambda and h-lengths in the upper half space model $\mathcal{U}$. A decorated ideal n-gon P is uniquely determined by a collection of pairwise disjoint points $s_i \in \hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, for $i = 1, \cdots, n$, together with a collection of Euclidean diameters δ_i of corresponding horocycles and perhaps the y-coordinate δ_∞ of a horocycle about ∞. Evidently these are actually coordinates, not just parameters, as they satisfy no relations other than pairwise inequality among the $\{s_i\}_1^n$.

The collection of pairs (s_i, δ_i), for $i = 1, \ldots, n$, are called *framed holographic coordinates* on the space of all decorated ideal polygons, and these are our desired coordinates. To explain the relationship with framed polygons, take a fixed copy of $\mathbb{D}$ and choose a fixed Cayley map $C^{-1} : \mathbb{D} \to \mathcal{U}$. Thus, given a $M\ddot{o}b$-orbit of decorated ideal polygon in $\mathbb{D}$ with framing $(u, v, w) \in C_3$, there is a well-defined decorated ideal polygon in $\mathcal{U}$ gotten by post-composing C with the unique Möbius transformation mapping $C^{-1}(u), C^{-1}(v), C^{-1}(w)$ to the respective points $\frac{0}{1}, \frac{1}{0}, \frac{1}{1}$.

We have proved

Theorem 2.5. *The moduli space of all framed decorated ideal n-gons modulo the action of Möb is given by the pairwise distinct n-tuples $\{s_i\}_1^n$ of centers of horocycles in $\hat{\mathbb{R}}$ with Euclidean diameters $\{\delta_i\}_1^n$ called the* framed holographic coordinates $\{(s_i, \delta_i) : 1 \leqslant i \leqslant n\}$. *The action of $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in$ Möb on these coordinates is the usual right fractional linear action $s_i \mapsto \frac{ds_i - b}{-cs_i + a}$ on the $\{s_i\}_1^n$, while the $\{\delta_i\}_1^n$ scale by the derivative of A at s, so $\delta_i \mapsto \frac{\delta_i}{(a-cs_i)^2}$.* □

Proof. The only point requiring clarification is the transformation law for δ coordinates, which follows from the fact [23, 26] that Euclidean diameters of horocycles scale by the derivative under a Möbius transformation. □

Remark 2.6. It is unsurprising that our framed holographic coordinates are reminiscent of the dihedral coordinates of [4] since the latter, which pertain only to the case of planar surfaces, are effectively related to the antecedent lambda lengths from [23] for any punctured surface, which are elaborated in the new coordinates.

Turning finally to stabilization given a framing $\mathcal{F} = (u, v, w) \in C_3$, define the matrix

$$L_{\mathcal{F}} = \frac{1}{(w-v)(u-w)(v-u)} \begin{pmatrix} v(u-w) & u(w-v) \\ u-w & w-v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$

which maps $(u, v, w) \mapsto (\frac{0}{1}, \frac{1}{0}, \frac{1}{1})$ under the right action, and given a second framing $\bar{\mathcal{F}}$, define

$$L_{\mathcal{F}}^{\bar{\mathcal{F}}} = L_{\mathcal{F}} \circ L_{\bar{\mathcal{F}}}^{-1} = \begin{pmatrix} a_{\mathcal{F}}^{\bar{\mathcal{F}}} & b_{\mathcal{F}}^{\bar{\mathcal{F}}} \\ c_{\mathcal{F}}^{\bar{\mathcal{F}}} & d_{\mathcal{F}}^{\bar{\mathcal{F}}} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}).$$

Now, given holographic coordinates $\{(s_i, \delta_i)\}_1^n$ with framing $\mathcal{F} = (u, v, w)$ and given (s, δ) with $s \notin \{s_i\}_1^n$, there is a unique $i^* \in \{1, \ldots, n\}$ with $s_{i^*} < s < s_{i^*+1}$. For $j = 1, \ldots, n+1$, define

$$(\bar{s}_j, \bar{\delta}_j) = \begin{cases} (s_j, \delta_j), & \text{if } j \leqslant i^*; \\ (s, \delta), & \text{if } j = i^* + 1; \\ (s_{j+1}, \delta_{j+1}), & \text{if } j > i^*, \end{cases}$$

and given a framing $\bar{\mathcal{F}} = (\bar{u}, \bar{v}, \bar{w})$ on $\{(\bar{s}_j, \bar{\delta}_j)\}_1^{n+1}$, finally let

$$\begin{aligned} s_k' &= (\bar{s}_k) L_{\mathcal{F}}^{\bar{\mathcal{F}}}, \\ \delta_k' &= (\bar{\delta}_k) \frac{d}{ds}\left((\bar{s}_k) L_{\mathcal{F}}^{\bar{\mathcal{F}}} \right), \end{aligned}$$

for $k = 1, \ldots, n+1$.

Corollary 2.7. *In the notation above, the mapping*

$$\{(s_i,\delta_i)\}_1^n \mapsto \{(s_i',\delta_i')\}_1^{n+1}$$

describes the stablization of framed holographic coordinates. In particular, if the framing $(u,v,w)=(\bar{u},\bar{v},\bar{w})$ *is constant, then the stablization is given by inclusion and re-indexing.* □

Remark 2.8. Furthermore according to [23, 26], the lambda length between (s,δ) and (s',δ') for $s,s'\neq\infty$ is given by $(\delta\delta')^{-\frac{1}{2}}|s-s'|$ and between (s,δ) and (∞,δ_∞) is given by $(\delta_\infty/\delta)^{\frac{1}{2}}$. On the level of the holographic parameters in Proposition 2.2 and in the foregoing notation, let c denote the lambda length of the edge between s_{i*} and s_{i*+1}, and let h_{i*} and h_{i*+1} denote the respective nearby h-lengths. If a,b denote the respective lambda lengths of the decorated edges between s and s_{i*}, s_{i*+1}, then after stabilization, the h-lengths at s_{i*}, s and s_{i*+1} are respectively given by $h_{i*}+\frac{b}{ac}$, $\frac{c}{ab}$ and $h_{i*+1}+\frac{a}{bc}$.

3. Framed and decorated homemorphisms of $\mathbb{S}^1$

Having developed framings on polygons as an alternative to the specification of a doe in the previous section, let us revisit from this framed point of view a basic result, Theorem 2.3 of [24], whose proof we first recall.

Theorem 3.1. *The space*

$$\mathcal{T}ess' = \{\text{tesselations of } \mathbb{D} \text{ with doe}\}$$

with topology induced by the Hausdorff topology on $\cup\tau\subset\mathbb{D}$ *is naturally homeomorphic to the space*

$$\mathrm{Homeo}_+ \approx \mathcal{T}ess'$$

with the compact-open topology.

Proof. Given $f\in\mathrm{Homeo}_+$, the image $\tau=f(\tau_*)$ of the Farey tesselation is another tesselation of $\mathbb{D}$, and the canonical doe $e_I\in\tau_*$ maps to the doe $e=f(e_I)$ of $\tau'=(\tau,e)\in\mathcal{T}ess'$ by definition. For the inverse map given a tesselation $\tau'=(\tau,e)$ with doe $e\in\tau$, begin by mapping the respective initial and terminal points of e_I to the initial and terminal points of e. Continue by mapping the further points in τ_*^0 of the respective triangles containing $|e_I|$ complementary to τ_* to the left and right of e_I in τ_* to those to the left and right of e in τ. Proceed in this way, essentially relying on the combinatorial rigidity of a tesselation with doe, in order to define $f^0 : \tau_*^0 \to \tau^0$, an order-preserving injection between dense subsets by construction. Using that a tesselation is locally-finite in $\mathbb{D}$ by definition, f^0 is seen to be a surjection as well, which thus interpolates a homeomorphism $f_{\tau'} \in \mathrm{Homeo}_+$ called the *characteristic map* of τ'. That the bijective assignment $\tau' \leftrightarrow f_{\tau'}$ is bicontinuous for the stated topologies is clear, completing the proof. ☐

A *framing on a tesselation* τ of $\mathbb{D}$ is the specification of an element of the configuration space C_3 of distinct triples in $\mathbb{S}^1$, whose members are not required to lie in τ^0, where the framed version of the basic result is

Corollary 3.2. *The space*

$$\mathcal{T}ess^f = \{\text{tesselations of } \mathbb{D}\}$$

is naturally homeomorphic to $\mathcal{T}ess' \approx \mathrm{Homeo}_+$*.*

Proof. As before for polygons, a doe e for τ determines an associated framing, and this gives an inclusion $\mathcal{T}ess' \subset \mathcal{T}ess^f$. Conversely given a framing (u, v, w) on the tesselation τ, choose a doe e_1 on τ with associated framing (u_1, v_1, w_1), let $L_1 = L_{u,v,w}^{u_1,v_1,w_1} \in M\ddot{o}b$ as in the previous section, and define

$$\tau'(\tau, u, v, w, e_1) = L_1^{-1} \circ f_{\tau_1'}(\tau_*) \in \mathcal{T}ess'$$

with its doe induced from e_I on τ_*.

We claim that $\tau_1' = \tau'(\tau, u, v, w, e_1)$ is independent of the choice of doe e_1 and to this end choose another doe e_2 on τ with associated framing (u_2, v_2, w_2), let $\tau_2' = \tau'(\tau, u, v, w, e_2)$ and $L_2 = L_{u,v,w}^{u_2,v_2,w_2}$. The composition $(L_2^{-1} \circ f_{\tau_2'})^{-1} \circ (L_1^{-1} \circ f_{\tau_1'})$ lies in PSL_2 by Lemma 1.1 since it leaves τ_* invariant, and in fact, it must be the identity since it furthermore pointwise fixes u, v, w. It follows that τ_1' and τ_2' have the same characteristic maps, and hence $\tau_1' = \tau_2'$ as tesselations with doe.

The assignment $(\tau, u, v, w) \mapsto L_1^{-1} \circ f_{\tau_1'}$ is thus a well-defined inverse to the inclusion $\mathcal{T}ess' \subseteq \mathcal{T}ess^f$ induced by associated framings, which is thus a homeomorphism as required. ☐

Let us emphasize that framings thus replace distinguished oriented edges in all regards: from determining global affine coordinates in this section, to the identification of $\mathcal{T}ess^f \approx \mathcal{T}ess'$ with Homeo_+, to the isomorphisms $\mathrm{Pt} \approx \mathrm{T} \approx \mathrm{PPSL}_2(\mathbb{Z})$ as well as their actions on tesselations.

It is straight-forward to describe the group structure on $\mathcal{T}ess^f$ induced by composition of homeomorphisms in Homeo_+ but difficult to visualize on the level of tesselations, except in special cases.

Proposition 3.3. *Consider framings $\mathcal{F}_X$ on the Farey tesselation τ_* associated to does e_X for $X \in PSL_2$. Then*

$$(\tau_*, \mathcal{F}_A) \circ (\tau_*, \mathcal{F}_B) = (\tau_*, \mathcal{F}_{BA}),$$

with the analogous statement holding for any tesselation. □

Next we introduce decorated versions of the foregoing spaces and begin by defining

$$\widetilde{\mathcal{T}ess^f} = \{\text{decorated and framed tesselations of } \mathbb{D}\},$$

where a *decoration* on a tesselation τ is the specification of one horocycle centered at each point of τ^0, so there is a natural forgetful mapping to $\widetilde{\mathcal{T}ess^f} \to \mathcal{T}ess^f$, whose fiber can be identified with $\mathbb{R}^\omega$.

Taking the quotient by the *Möb*-action on framings, we have

$$\begin{aligned} \mathcal{T}ess^f/M\ddot{o}b \approx \mathcal{T}ess^n &= \left\{\text{tesselations } \tau : \tfrac{0}{1}, \tfrac{1}{1}, \tfrac{1}{0} \in \tau^0\right\}, \\ \widetilde{\mathcal{T}ess^f}/M\ddot{o}b \approx \widetilde{\mathcal{T}ess^n} &= \left\{\text{decorated tesselations } \tau : \tfrac{0}{1}, \tfrac{1}{1}, \tfrac{1}{0} \in \tau^0\right\} \end{aligned}$$

where the superscript n stands for *normalized.* Notice that the Farey tesselation τ_* is itself already normalized.

As for the analogous elaboration of circle homemorphisms, first define spaces

$$\text{Homeo}_+^n = \{f \in \text{Homeo}_+ : f(t) = t \text{ for } t = \tfrac{0}{1}, \tfrac{1}{1}, \tfrac{1}{0}\}$$

of *normalized* homeomorphisms, so $\mathcal{T}ess^n \approx \text{Homeo}_+^n$ as Homeo_+^n-spaces. Next define the decorated version

$$\widetilde{\text{Homeo}_+^n} = \{(\tilde{f}, f) : f \in \text{Homeo}_+^n \text{ covered by } \tilde{f} : \widetilde{\mathcal{T}ess^n} \to \widetilde{\mathcal{T}ess^n}\}$$

with the natural group structure, making $(\tilde{f}, f) \mapsto f$ a group homomorphism $\widetilde{\text{Homeo}_+^n} \to \text{Homeo}_+^n$. There is analogously an isomorphism $\widetilde{\text{Homeo}_+^n} \approx \widetilde{\mathcal{T}ess^n}$ of $\widetilde{\text{Homeo}_+^n}$-spaces, where (identity,identity) corresponds to the Farey tesselation with its canonical decoration $\tilde{\tau}_*$, namely, the one determined by the condition that the horocycles at endpoints of any common geodesic are taken to be osculating.

This leads to the commutative cube

$$\begin{tikzcd}
\widetilde{\mathcal{T}ess^f} \arrow[rr, "\approx"] \arrow[dd] \arrow[dr] & & \widetilde{\mathrm{Homeo}_+} \arrow[dd] \arrow[dr] & \\
& \mathcal{T}ess^f \arrow[rr, "\approx"] \arrow[dd] & & \mathrm{Homeo}_+ \arrow[dd] \\
\widetilde{\mathcal{T}ess^n} \arrow[rr, "\approx"] \arrow[dr] & & \widetilde{\mathrm{Homeo}^n_+} \arrow[dr] & \\
& \mathcal{T}ess^n \arrow[rr, "\approx"] & & \mathrm{Homeo}^n_+
\end{tikzcd}$$

where $\widetilde{\mathrm{Homeo}_+}$ is defined by pull-back, the vertical maps are principal $M\ddot{o}b$-bundles, and the maps out of the plane of the page are $\mathbb{R}^\omega$-bundles.

Theorem 3.4. *Lambda lengths on edges of* τ_* *give global affine coordinates on* $\widetilde{\mathcal{T}ess^n} \subseteq \mathbb{R}_+^{\tau_*}$. *Indeed, there is an explicit construction of a Möb-orbit of tesselations of* $\mathbb{D}$ *with doe from the assignment of a putative lambda length to each edge of* τ_*. $\square$

The proof of this result as Theorem 3.1 in [24] provides the explicit recursive construction of a $M\ddot{o}b$-orbit of decorated tesselations of $\mathbb{D}$ with doe from the assignment of a putative lambda length to each edge of τ_*, in analogy to the proof of Theorem 2.1. The canonical decoration $\tilde{\tau}_*$ on τ_*, or in other words the identity element of the group $\widetilde{\mathrm{Homeo}^n_+}$, corresponds to taking all these lambda length coordinates equal to unity. In the classical setting of a suitably decorated punctured surface of finite topological type uniformized by a torsion free subgroup of finite-index in PSL_2, or a so-called *punctured arithmetic surface*, the classical lambda length coordinates on decorated Teichmüller space from [23] are likewise all unity, so this identity element in $\widetilde{\mathrm{Homeo}^n_+}$ may be regarded as the universal punctured arithmetic surface.

Remark 3.5. It is not difficult to descend these lambda length coordinates on the bottom-back of the previous commutative cube to its bottom-front by assigning cross ratios $\frac{ac}{bd}$, or shear coordinates $\ln\frac{ac}{bd}$, instead of lambda lengths e to each edge, where e is the diagonal of the quadrilateral with frontier edges a, b, c, d in this cyclic order; see [24].

We next discuss framed tesselations as limits of framed polygons. To this end, notice that there is the natural linear ordering on $\hat{\mathbb{Q}}$ arising from the lexicographic ordering on pairs given by Farey generation followed by the counter-clockwise order in $\mathbb{S}^1$ starting from $\frac{0}{1}$, and there is thus an induced bijective enumeration $\mathbb{Z}_{\geqslant 0} \to \hat{\mathbb{Q}}$.

Remark 3.6. This is the raison d'être for the Farey construction. Indeed, the mineralogist Farey published without proof this solution to the long-standing open problem of producing an explicit bijective enumeration of the rational numbers, and the proof was essentially immediately supplied by Cauchy.

Construction 3.7. Suppose that $S \subseteq \mathbb{S}^1$ is a countable dense subset enumerated by the bijection $\mu : \mathbb{Z}_{\geqslant 0} \to S$. Construct a function $f_\mu : \hat{\mathbb{Q}} \to S$ as follows. For the basis step of our recursive construction, define $f_\mu(\frac{0}{1}) = \mu(0)$ and $f_\mu(\frac{1}{0}) = \mu(1) \in \mathbb{S}^1$. Recursively suppose that the images of the Farey points of generation at most $g \geqslant 0$ have been defined. This collection of image points decomposes $\mathbb{S}^1$ into 2^{g+1} open circular intervals. Each such interval with endpoints $x = f_\mu(\frac{p}{q}), y = f_\mu(\frac{r}{s})$ contains a point $z \in S$ of least index $\mu^{-1}(z)$, and we define $f_\mu(\frac{p+r}{q+s}) = z$ in this case, thereby extending the function f_μ to generation $g + 1$. We furthermore derive a collection $\tau_\mu = \{f_\mu(e) : e \in \tau_*\}$ of geodesics with doe $f_\mu(e_I)$, where $f_\mu(e)$ denotes the geodesic in $\mathbb{D}$ with endpoints the f_μ-image of the endpoints of $e \in \tau_*$.

Suppose that $\mu : \mathbb{Z}_{\geqslant 0} \to S$ is a bijective enumeration of a countable dense subset $S \subseteq \mathbb{S}^1$. A circular interval $I \subseteq \mathbb{S}^1$ with endpoints $x, y \in S$ is said to be *solid* provided $\mu^{-1}(z) > \max\{\mu^{-1}(x), \mu^{-1}(y)\}$ for every z in the interior of I. The enumeration μ is said to be *convergent* if every infinite proper nested family $I_0 \supsetneq I_1 \supsetneq \cdots$ of solid open intervals is disjoint from S, that is, $S \cap \bigcap_{j \geqslant 0} I_j = \varnothing$.

For any $e \in \tau_*$ other than the doe, exactly one of the circular intervals in $\mathbb{S}^1$ complementary to its endpoints is solid, and the Farey enumeration of $\hat{\mathbb{Q}}$ is convergent by definition of the Farey ordering. We may always assume that an enumeration of a countable dense subset is indexed by the Farey tesselation in its canonical linear ordering.

Proposition 3.8. *For any convergent bijective enumeration $\hat{\mathbb{Q}} \to S$ of a countable dense subset $S \subset \mathbb{S}^1$, Construction 3.7 yields a bijection $f_\mu : \hat{\mathbb{Q}} \to S$ and a tesselation $\tau_\mu = f_\mu(\tau_*)$ with doe. Conversely, if $f_\mu : \hat{\mathbb{Q}} \to S$ is a bijection, then τ_μ is a tesselation, and μ must be convergent.*

Proof. Suppose that μ is convergent and $z \in S$. For $g \geqslant 0$, let I_g denote the component of $\mathbb{S}^1 - \{f_\mu(x) : x \in \hat{\mathbb{Q}} \text{ is of generation at most } g\}$ which contains z. It follows that $I_0 \supsetneq I_1 \supsetneq \cdots$ is a nested family of solid intervals. This sequence must terminate, for otherwise $z \in S \cap \bigcap_{g \geqslant 0} I_g \neq \varnothing$ contradicts convergence of μ. Thus, f_μ maps onto S and is injective by construction, and so is a bijection.

To see that $\tau_\mu = f_\mu(\tau_*)$ is locally finite, suppose in order to derive a contradiction that a sequence of points in $e_i = f_\mu(e_i')$, for $i \geqslant 1$ and $e_i' \in \tau_*$, accumulates at some point in $\mathbb{D}$. It follows that e_i limits to some geodesic e_∞ in $\mathbb{D}$. Since the $\{e_i\}$ are pairwise disjoint in $\mathbb{D}$ by construction, we may assume that they all lie on one side of e_∞. By density of S in $\mathbb{S}^1$, there is some $z \in S$ on the other side of e_∞, and it lies in the f_μ-image of some generation g point of $\hat{\mathbb{Q}}$ by the already established surjectivity. Taking a further subsequence if necessary, the complementary intervals to the endpoints of e_i that contains z are solid. The generation of these endpoints is therefore bounded above by g, so that $\{e_i\}$ is a finite set, as required. The same argument shows that each component of $\mathbb{D} - \cup\tau_\mu$ is an ideal triangle, so τ_μ is indeed a tesselation.

Conversely, suppose that $f_\mu : \mathbb{Q} \to S \subseteq \mathbb{S}^1$ is surjective and $z \in S \cap \bigcap_{j \geqslant 1} I_j$ for some nested sequence $I_0 \supsetneq I_1 \supsetneq \cdots$ of solid intervals with endpoints in S. The point $z \in S$ must a fortiori have some fixed generation, which bounds above the generations of the endpoints of the intervals. Since there are only finitely many points of any given generation, the sequence must terminate, as required. $\square$

It is an easy matter now to stabilize nested, decorated and framed ideal polygons:

Theorem 3.9. *Consider a nested sequence*

$$P_0 \subseteq P_1 \subseteq \cdots \subseteq P_g \subseteq P_{g+1} \subseteq \cdots$$

of ideal polygons with a common framing $\mathcal{F} \in C_3$, where P_0 is a geodesic, P_g has 2^{g+1} sides for $g \geqslant 1$, and $P_{g+1} - P_g$ consists of 2^g ideal triangles, for all $g \geqslant 0$. If the union of the ideal vertices of all the polygons is dense, then there is a well-defined limiting tesselation also framed by $\mathcal{F}$, and any such is conversely given by such a limit. The analogous statement holds for decorated framed polygons and decorated framed tesselations. Moreover, the tuple of framed holographic coordinates

$$(s_{\frac{0}{1}}, \delta_{\frac{0}{1}}), (s_{\frac{1}{0}}, \delta_{\frac{1}{0}}), (s_{\frac{1}{1}}, \delta_{\frac{1}{1}})(s_{-\frac{1}{1}}, \delta_{-\frac{1}{1}}), \ldots$$

in $(\hat{\mathbb{R}} \times \mathbb{R}_{>0})^\omega$ *of decorated polygons in the weak topology given in their natural Farey ordering, where the* $\{s_i\}$ *are required to be pairwise distinct and* $\cup\{s_i\}$ *to be dense in* $\hat{\mathbb{R}}$, *provide global coordinates on* $\widetilde{\mathcal{T}ess^f} \approx \widetilde{\mathrm{Homeo}}_+$. $\square$

4. The Lie algebra *ppsl*$_2$

Let $sl_2 = sl_2(\mathbb{R})$ denote the usual Lie algebra of traceless two-by-two real matrices with generators $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $f = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and $h = \left(\begin{smallmatrix} +1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and Lie brackets $[h, e] = 2e$, $[h, f] = -2f$ and $[e, f] = h$. (Context will distinguish the notational inconvenience of determining between $e \in sl_2$ or $e \in \tilde{\tau}$.)

Exponentiating $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & -\alpha \end{smallmatrix}\right) \in sl_2$ yields a one-parameter family of diffeomorphisms giving rise to a vector field on the circle which is given by

$$\{(\gamma + \beta) \cos \theta + 2\alpha \sin \theta + (\gamma - \beta)\} \frac{\partial}{\partial \theta},$$

where $\frac{\partial}{\partial \theta}$ denotes the constant unit vector field on $\mathbb{S}^1$ and θ its usual angular coordinate. A vector field A on $\mathbb{S}^1$ arising in this way is called a *(global) sl_2 vector field*, and we write $A \in sl_2$ in this case. The values of an sl_2 vector field at any three distinct points in $\mathbb{S}^1$ determine it uniquely. In case two vector fields ϑ_1 and ϑ_2 on $\mathbb{S}^1$ differ by a global sl_2 vector field, then we shall write $\vartheta_1 \doteq \vartheta_2$.

More generally, a vector field ϑ on $\mathbb{S}^1$ is a *piecewise sl_2 vector field* if $\mathbb{S}^1$ decomposes into *finitely many* open connected circular intervals with pairwise disjoint interiors whose endpoints are among the rational points of $\mathbb{S}^1$ so that ϑ restricts on the interior of each such interval to some global sl_2 vector field. $ppsl_2$ denotes the collection of all such vector fields including the possibility of no breakpoints, namely $sl_2 \subseteq ppsl_2$ itself. The endpoints of the *maximal such intervals* are called the *breakpoints* of ϑ itself. There are no restrictions on the behavior of ϑ at its breakpoints (except that the breakpoints, if any, are rational), and indeed $\vartheta \in ppsl_2$ may not even be defined at its breakpoints in general, in which case its value is implicitly given as the average of the two one-sided limits.

Remark 4.1. Let us parenthetically recall a seminal result of Dirichlet, which he proved in Berlin at age 24 thereby besting the top mathematicians of his day: the Fourier series of a piecewise smooth function with finitely many pieces converges pointwise to the function itself except at the breakpoints where it is the average of its two one-sided limits.

A bracket $[\vartheta_1, \vartheta_2] \in ppsl_2$ of two $\vartheta_1, \vartheta_2 \in ppsl_2$ is defined in the natural way: the resulting vector field has preliminary breakpoints given by the union of those of ϑ_1 and those of ϑ_2, and on each complementary component in $\mathbb{S}^1$ of this union, the bracket is given by the usual bracket on sl_2. It may happen that the actual breakpoints of $[\vartheta_1, \vartheta_2]$ form a proper subset of the preliminary breakpoints since we demand maximality of intervals complementary to the actual breakpoints.

Consider breakpoints $\pm 1, \pm i \in \mathbb{S}^1 \subseteq \mathbb{C}$, so the complementary intervals lie in respective quadrants I-IV in the complex plane enumerated as usual in the counterclockwise sense beginning with quadrant I where both coordinates are non-negative, and define

$$\Lambda(s) = \begin{cases} \begin{pmatrix} s & s-s^{-1} \\ 0 & s^{-1} \end{pmatrix}, & \text{on quadrant I;} \\ \begin{pmatrix} s^{-1} & 0 \\ s-s^{-1} & s \end{pmatrix}, & \text{on quadrant II;} \\ \begin{pmatrix} s^{-1} & 0 \\ s^{-1}-s & s \end{pmatrix}, & \text{on quadrant III;} \\ \begin{pmatrix} s & s^{-1}-s \\ 0 & s^{-1} \end{pmatrix}, & \text{on quadrant IV,} \end{cases}$$

for $s \in \mathbb{R} - \{0\}$, a one-parameter family in $\mathrm{PPSL}_2(\mathbb{R})$. It is not difficult to check that each $\Lambda(s)$ is moreover once-continuously differentiable on $\mathbb{S}^1$ including at its breakpoints.

Since lambda lengths are *Möb*-invariant, it is clear that $\Lambda(s)$ leaves invariant all lambda lengths on τ_* with any decoration, except the lambda length of the doe, which it scales by a factor s. Thus, $\Lambda(s)$ is a one-parameter family in $\mathrm{PPSL}_2(\mathbb{R})$ scaling just one affine lambda length coordinate, or in other words a multiplicative coordinate deformation.

More specifically, $\Lambda(s) \in \mathrm{PPSL}_2(\mathbb{R})$ is a one-parameter family which on each circular interval determined by the intersection of a quadrant with $\mathbb{S}^1$ is a hyperbolic transformation whose axis is spanned by the endpoints of the interval. In the parlance of the Thurston school, an *earthquake* is an element of $\mathrm{PPSL}_2(\mathbb{R})$ with two pieces, one of which is the identity and the other is a hyperbolic transformation whose axis is asymptotic to the breakpoints. Thus, $\Lambda(s)$ is a special one-parameter family of compositions of four earthquakes, where the four hyperbolic translations are chosen in order to produce homeomorphisms $\Lambda(s)$ that are once-continuously differentiable on $\mathbb{S}^1$.

The derivative of $\Lambda(s)$ with respect to s at $s = 1$ is the extremely special element $\vartheta \in ppsl_2$ illustrated in Figure 4 and called the *mother* *wavelet*. It is the basic building block of $ppsl_2$. Justification for the appellate "wavelet" used here is given in [27].

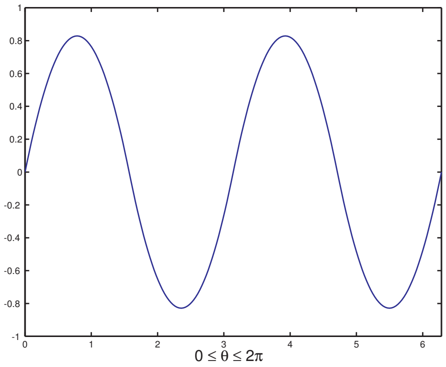


Figure 4. The mother wavelet ϑ.

Direct computation confirms that the mother wavelet ϑ is given by

$$\vartheta = \begin{cases} +h + 2e, & \text{in quadrant I;} \\ -h + 2f, & \text{in quadrant II;} \\ -h - 2f, & \text{in quadrant III;} \\ +h - 2e, & \text{in quadrant IV} \end{cases}$$

and vanishes at each of its breakpoints, where it is once-continuously differentiable. Notice that in general if an element of $ppsl_2$ is *twice* continuously differentiable at a breakpoint, then it is not actually a breakpoint at all, that is, if two elements of $M\ddot{o}b$ agree to second order at a point, then they must coincide.

Let us employ the adjoint action on each piece and define the *(arithmetic) wavelets*

$$\vartheta_A(\theta) = A^{-1}\vartheta_A(\theta.A)A, \text{ for } A \in \mathrm{PSL}_2,$$

where the right A-action on $\theta \in \mathbb{S}^1$ is the natural one. A short calculation shows that $\vartheta_S = \vartheta_I = \vartheta$, and therefore if A_+, A_- correspond to the two different orientations on a common edge, i.e., if $A_\mp = SA_\pm$ then $\vartheta_{A_+} = \vartheta_{A_-}$. It follows that for any unoriented edge we have $|e_A| = |e_{SA}|$, and there is a corresponding vector field $\vartheta_{|e_A|} = \vartheta_A = \vartheta_{SA}$.

Now for each $A \in \mathrm{PSL}_2$, define a corresponding $X_A \in sl_2$ where X_A and ϑ_A take the same values at the three points $\pm 1, -i \in \mathbb{S}^1$, or in other words at Farey points $\frac{0}{1}, \frac{1}{0}, \frac{1}{1}$, and define the *normalization*

$$\bar{\vartheta}_A = \vartheta_A - X_A, \text{ for } A \in \mathrm{PSL}_2.$$

It is not difficult to compute X_A explicitly for $A \in \mathrm{PSL}_2$.

Here and herein we shall deviate slightly from the notation of previous works, which we first recall.

Following [20, 26] for each $A \in \mathrm{PSL}_2$, define for each $A \in \mathrm{PSL}_2$ the infinite sums

$$\bar{\phi}_A = \sum_{n \geqslant 0} \bar{\vartheta}_{U^n A} \text{ and } \bar{\phi}^*_A = \sum_{n \leqslant 0} \bar{\vartheta}_{U^n A}$$

respectively called *normalized left and right fans* and the further infinite sums

$$\bar{\psi}_A = \sum_{n \geqslant 0} n \; \bar{\vartheta}_{U^n A} \text{ and } \bar{\psi}^*_A = \sum_{n \leqslant 0} n \; \bar{\vartheta}_{U^n A}$$

respectively called *normalized left* and *right hyperfans*. We have added bars to the notation for normalized (hyper)fans from earlier work and keep the undecorated symbol for more natural normalizations to be introduced in the next section. It is not difficult to compute that $\sum_{n \geqslant 0} X_{U^n A}$ diverges thus explaining the need for normalization.

Assuming for the moment that these sums converge, as we shall discuss presently, there is the following prescribed consequence of this "hyperfan formalism" given by infinite sums of infinite sums in this way.

Proposition 4.2. *For each $A \in \mathrm{PSL}_2$, we have*

$$\bar{\psi}_A = \sum_{n \geqslant 1} \bar{\phi}_{U^n A} \text{ and } \bar{\psi}^*_A = \sum_{n < 0} \bar{\phi}^*_{U^n A},$$

as well as

$$\begin{aligned} \bar{\psi}_A - \bar{\psi}_{UA} &= \bar{\phi}_{UA} \text{ and } \bar{\phi}_A - \bar{\phi}_{UA} = \vartheta_A, \\ \bar{\psi}^*_{UA} - \bar{\psi}^*_A &= \bar{\phi}^*_A \text{ and } \bar{\phi}^*_{UA} - \bar{\phi}^*_A = \vartheta^*_A \end{aligned}$$

□

Corollary 4.3. *For each $A \in \mathrm{PSL}_2$, equating $\bar{\vartheta}_A = \bar{\vartheta}_{SA}$ gives the so-called* USA relation

$$\begin{aligned} \bar{\psi}_{UA} - 2\bar{\psi}_A + \bar{\psi}_{U^{-1}A} &= \bar{\psi}_{USA} - 2\bar{\psi}_{SA} + \bar{\psi}_{U^{-1}SA}, \\ \bar{\psi}^*_{UA} - 2\bar{\psi}^*_A + \bar{\psi}^*_{U^{-1}A} &= \bar{\psi}^*_{USA} - 2\bar{\psi}^*_{SA} + \bar{\psi}^*_{U^{-1}SA}. \end{aligned}$$

□

The formulas in Proposition 4.2 and Corollary 4.3 thus follow immediately from the hyperfan formalism and the symmetry $\bar{\vartheta}_A = \bar{\vartheta}_{SA}$. The catch is showing that the putative (hyper)fans converge. However, notice that convergence of $\sum n x_n$ implies that of $\sum x_n$, so the hyperfan formalism follows from convergence of hyperfans alone.

In the remainder of this section, we shall recall results from [20] which in particular give sense to the *normalized* (hyper)fans as elements of $ppsl_2$, and indeed converge pointwise uniformly on compacta. The harmonic analysis of normalized hyperfans is discussed in Appendix A.

The **first big surprise** is that in fact normalized fans and hyperfans have only finitely many breakpoints and lie in $ppsl_2$, even though they are defined by infinite sums. Indeed, a fan always has exactly three breakpoints and is described by a continuous but not differentiable function on the circle, and a hyperfan always has exactly two breakpoints described by a function on the circle which is discontinuous at exactly one point. Such is the nature of telescoping for normalized (hyper)fans, which are *not enjoyed* by unnormalized (hyper)fans.

For example, $\bar{\phi}_U$ takes values $-2e$ on quadrant I, $2(h-f)$ on quadrant II and vanishes on quadrants III and IV, while $\bar{\psi}_I$ takes values $-2e$ on quadrants I and II and vanishes on quadrants III and IV. In fact, one finds that $\bar{\psi}_I + \bar{\psi}_I^*$ is the global sl_2 vector field $-2e$, and together with the fact that $\psi_A \doteq A^{-1}\psi_I A$ by construction, it follows that the span of the left hyperfans together with sl_2 contains the right hyperfans. We shall therefore henceforth restrict our attention to the former and drop the appellate "left" tacitly taking only left fans and hyperfans. There is an entirely parallel discussion using right fans and hyperfans.

Just as one might suspect, bracketing destroys one degree of smoothness. Brackets of normalized wavelets are similarly expressed as finite sums of normalized fans while brackets of normalized fans are expressed in terms of normalized hyperfans.

The **second big surprise** is that here the algebra closes with the additive basis of normalized left hyperfans, i.e., brackets of hyperfans are finite linear combinations of hyperfans.

Summarizing several of the main achievements in [20], to which we refer the reader for proofs of the first two big surprises, we have

Theorem 4.4. *The set of normalized left hyperfans together with the generators $e, f, h \in sl_2$ give an additive spanning set for the vector space $ppsl_2$ which is closed under bracketing. Moreover, the collection of USA relations in Corollary 4.3, one for each edge of τ_*, gives a complete set of relations among normalized left hyperfans.* □

In fact, there is an error in the proof in [20] that normalized left hyperfans span $ppsl_2$ which is corrected in Theorem 2.1 of [27] as follows.

Theorem 4.5. *Define orientations on the edges of τ_* with the orientation from $\frac{0}{1}$ to $\frac{1}{0}$ as usual and otherwise always pointing from lower to* *higher generation in the Farey enumeration, and denote this set of oriented edges* $\mathcal{O} \subset \tilde{\tau}_*$. *Then* $\mathcal{B} = \{\bar{\psi}_A : A \in \mathcal{O}\}$ *together with* $e, f, h \in sl_2$ *is an additive basis for* $ppsl_2$. $\square$

5. The new formalism for $ppsl_2$

The normalization of vector fields in the previous section was necessary in order to guarantee their pointwise convergence and assure the hyperfan formalism. However from the point of view of representation theory, the failure of exact equality in favor of equality $\bar{\psi}_A \doteq A^{-1}\bar{\psi}_I A$ up to global sl_2 vector fields introduces untoward complications. Furthermore from the point of view of physics, the treatment of sl_2 in the previous section represents merely an additional copy of sl_2 essentially decoupled from the higher Fourier modes, not the creation/annihilation and energy operators one might anticipate.

Let us remedy both of these deficiencies by defining *(unnormalized)* (or really differently normalized) *hyperfans* as follows:

$$\psi_I = \begin{cases} e, & \text{on quadrants I and II;} \\ 0, & \text{on quadrants III and IV,} \end{cases}$$

thus dropping the pre-factor -2 from before and furthermore guaranteeing the desired conjugacy formula by defining

$$\psi_A(\theta) = A^{-1}\psi_I(\theta.A)A, \text{ for } A \in \mathrm{PSL}_2.$$

It thus follows that $B^{-1}\psi_A B = \psi_{AB}$.

The **third big surprise**, which is new to this paper, is that sl_2 is actually in the finitely supported span of these new hyperfans, as we next prove.

Proposition 5.1. *For* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}_2$, *define*

$$\Psi_A = \psi_{STA} - 2\psi_{SA} + \psi_{ST^{-1}A} - \{\psi_{UA} - 2\psi_A + \psi_{U^{-1}A}\}.$$

Then we have the identity

$$\Psi_A = \{c(d+b) + a(d-b)\}h \; + \; (d^2 - b^2 + 2bd)e \; + \; (a^2 - c^2 - 2ac)f.$$

In particular

$$\Psi_I = h + e + f, \;\; \Psi_T = 2e + f, \;\; \Psi_{U^{-1}} = e + 2f,$$

whence

$$\begin{pmatrix} h \\ e \\ f \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 & -1 & -1 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} \Psi_I \\ \Psi_T \\ \Psi_{U^{-1}} \end{pmatrix}$$

Proof. The reader will recognize Ψ_A as the difference of the two sides of the USA relation which holds for normalized hyperfans and fails by a global sl_2 for unnormalizaed hyperfans. The proposition follows from direct computation using that $\gamma - 2\beta = 2\alpha - \delta$, where

$$\begin{aligned} \alpha &= A^{-1}eA = cdh + d^2e - c^2f, \ \beta = (SA)^{-1}eSA = abh + b^2e - a^2f, \\ \gamma &= (UA)^{-1}e(UA) = (c+a)(d+b)h + (d+b)^2e - (c+a)^2f \\ &= (STA)^{-1}e(STA), \\ \delta &= (U^{-1}A)^{-1}e(U^{-1}A) = (c-a)(d-b)h + (d-b)^2e - (c-a)^2f \\ &= (ST^{-1}A)^{-1}e(ST^{-1}A). \end{aligned}$$

□

We shall refer to the expression for Ψ_A in terms of e, f, h given in Proposition 5.1 as the *USA identity* for $A \in \mathrm{PSL}_2$.

Theorem 5.2. *The hyperfans $\{\psi_A : A \in \mathrm{PSL}_2\}$ span $ppsl_2$, a complete set of relations on them is given by the USA identities as A varies over* PSL_2*, and $\{\psi_A : A \in \mathcal{O}\}$ provides a basis for $ppsl_2$, where $\mathcal{O} \subset \tilde{\tau}$ is given in Theorem 4.5.*

Proof. In the old formalism, the span of normalized hyperfans together with e, f, h was shown to be all of $ppsl_2$. This contains all unnormalized hyperfans by definition; indeed, $\psi_A \doteq \bar{\psi}_A$, and the reverse inclusion follows from the last part of the previous proposition. For the second assertion, suppose that $\sum \alpha_A \psi_A = 0$ is a finite linear relation among the unnormalized hyperfans. It follows that then $\sum \alpha_A \bar{\psi}_A \doteq 0$, and since both sides of this equation vanish at $\frac{0}{1}, \frac{1}{0}, \frac{1}{1}$, it must be that in fact $\sum \alpha_A \bar{\psi}_A = 0$, whence this relation is a consequence of the USA relations by Theorem 4.4. A non-trivial finite linear relation among $\{\psi_A : A \in \mathcal{O}\}$ likewise gives such a relation on $\{\bar{\psi}_A : A \in \mathcal{O}\}$, which would contradict Theorem 4.5. □

Another favorable aspect of the new formalism is that the structure constants of $ppsl_2$ admit an explicit if not entirely trivial expression, and it is to this end that we dedicate the remainder of this section. The next two results follow from direct and elementary computation, which are left to the reader.

Proposition 5.3. *In the basis e, f, h for sl_2, the adjoint $x \mapsto A^{-1}xA$ of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in sl_2$ is given by the matrix*

$$M_A = \begin{pmatrix} d^2 & -b^2 & 2bd \\ -c^2 & a^2 & -2ac \\ cd & -ab & ad+bc \end{pmatrix},$$

and so

$$M_A^{-1} = M_{A^{-1}} = \begin{pmatrix} a^2 & -b^2 & -2ab \\ -c^2 & d^2 & 2cd \\ -ac & bd & ad+bc \end{pmatrix}.$$

□

Just as ψ_I takes values $e \in sl_2$ on quadrants I and II and vanishes on quadrants III and IV, we have the following analogous result for $f, h \in sl_2$.

Proposition 5.4. *$\psi_S + f$ and $\psi_I + \psi_{US} - \psi_S - \psi_{U^{-1}} - f$ each have support on quadrants I and II with the former taking there the value f and the latter the value h.* □

Now in order to compute the bracket $[\psi_B, \psi_A]$, for $A, B \in \mathrm{PSL}_2$, it suffices to compute simply $[\psi_I, \psi_A]$, since we have

$$(*) \qquad [\psi_B, \psi_A] = [B^{-1}\psi_I B, \psi_A] = B^{-1}[\psi_I, \psi_{AB^{-1}}]B.$$

To this end with $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, there are four essential cases:

Case 1: $0 \leqslant -\frac{d}{c} < -\frac{b}{a}$, so e_A lies in the bottom half plane oriented from right to left.

Case 2: $0 \leqslant -\frac{b}{a} < -\frac{d}{c}$, so e_A lies in the bottom half plane oriented from left to right.

Case 3: $-\frac{d}{c} < -\frac{b}{a} \leqslant 0$, so e_A lies in the top half plane oriented from left to right.

Case 4: $-\frac{b}{a} < -\frac{d}{c} \leqslant 0$, so e_A lies in the top half plane oriented from right to left.

In each case, the support of ψ_A lies to the left of e_A by definition. In Case 1, the supports are therefore disjoint, and so $[\psi_I, \psi_A] = 0$. In Case 2, the support of ψ_I is contained in the support of ψ_A, and the bracket is supported on quadrants I and II taking value $-c^2h + 2cde$, which is given according to Proposition 5.4 by

$$\begin{aligned}-c^2(\psi_I + \psi_{US} - \psi_S - \psi_{U^{-1}} - f) + 2cd\psi_I \\ = c(2d-c)\psi_I + c^2(\psi_S + \psi_{U^{-1}} - \psi_{US} + f).\end{aligned}$$

Case 3 is more challenging requiring Proposition 5.3. We compute

$$M_A^{-1}\begin{pmatrix}2cd\\0\\-c^2\end{pmatrix} = \begin{pmatrix}2ac(ad+bc)\\-4c^3d\\-c^2(3ad+bc)\end{pmatrix},$$

whence

$$\begin{aligned}[\psi_I, \psi_A] &= 2ac(ad+bc)\psi_A - 4c^3d(\psi_{SA} + A^{-1}fA)\\ &\quad - c^2(3ad+bc)\{\psi_A + \psi_{USA} - \psi_{SA} - \psi_{U^{-1}} - A^{-1}fA\}.\end{aligned}$$

Case 4 is still a bit more involved since $[\psi_I, \psi_A]$ perhaps has two components and can be expressed as the difference $\mu-\nu$, where μ is supported on quadrants I and II taking value $2cde-c^2h$, and ν is supported on the region to the left of e_{SA} and taking the same value there. Meanwhile, Proposition 5.4 gives the expression

$$\mu = c(2d-c)\psi_I + c^2(\psi_S + \psi_{U^{-1}} - \psi_{US} + f)$$

as in Case 2. For the other term ν, we compute

$$\begin{aligned}M_{SA}^{-1}\begin{pmatrix}2cd\\0\\-c^2\end{pmatrix} &= \begin{pmatrix}c^2 & -d^2 & -2cd\\-a^2 & b^2 & 2ab\\ac & -bd & -(ad+bc)\end{pmatrix}\begin{pmatrix}2cd\\0\\-c^2\end{pmatrix}\\ &= \begin{pmatrix}4c^3d\\-2ac(ad+bc)\\c^2(3ad+bc)\end{pmatrix},\end{aligned}$$

whence

$$\begin{aligned}\nu =& 4c^3d\psi_{SA} - 2ac(ad+bc)(\psi_A + (SA)^{-1}fSA)\\ &+ c^2(3ad+bc)\{\psi_{SA} + \psi_{UA} - \psi_A - \psi_{U^{-1}SA} - (SA)^{-1}fSA\}.\end{aligned}$$

Putting all this together, a little further computation finally proves

Theorem 5.5. *For* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}_2$, *we have*

$$[\psi_I, \psi_A] = \begin{cases} 0, \text{ if } 0 \leqslant -\frac{d}{c} < -\frac{b}{a}; \\[2ex] c(2d-c)\psi_I + c^2(\psi_S + \psi_{U^{-1}} - \psi_{US} + f), \\ \quad \text{if } 0 \leqslant -\frac{b}{a} < -\frac{d}{c}; \\[2ex] 2ac(ad+bc)\psi_A - 4c^3 d(\psi_{SA} + A^{-1} f A) \\ -\, c^2(3ad+bc)\{\psi_A + \psi_{USA} - \psi_{SA} - \psi_{U^{-1}} - A^{-1} f A\}, \\ \quad \text{if } -\frac{d}{c} < -\frac{b}{a} \leqslant 0; \\[2ex] c(2d-c)\psi_I + c^2(\psi_S + \psi_{U^{-1}} - \psi_{US} + f) \\ -\, 4c^3 d\psi_{SA} + 2ac(ad+bc)(\psi_A + (SA)^{-1} f AS) \\ -\, c^2(3ad+bc)\{\psi_{SA} + \psi_{UA} - \psi_A - \psi_{U^{-1}SA} - (SA)^{-1} f AS\}, \\ \quad \text{if } -\frac{b}{a} < -\frac{d}{c} \leqslant 0. \end{cases}$$

□

It is worth noting that $[\psi_I, \psi_A] = 0$ if $c = 0$ and that to complete the calculation of brackets in the basis $\{\psi_A : A \in \mathcal{O}\}$, one must still conjugate as in equation (*) and finally use the constructive proof of our Theorem 5.2 given in [20] to express several of the resulting hyperfans in terms of this basis. The point, however, is that brackets are explicitly computable in the new formalism in contrast to the old.

6. Three 2-forms

As before, let $\frac{\partial}{\partial\theta}$ denote the counter-clockwise unit vector field on $\mathbb{S}^1$ in its usual angular coordinate θ and $L_n = ie^{in\theta}\frac{\partial}{\partial\theta}$, for $n \in \mathbb{Z}$, denote the usual generators for the *Witt algebra*, satisfying for $m, n \in \mathbb{Z}$ the bracket identity $[L_m, L_n] = (m-n)\ L_{m+n}$. The Witt algebra is naturally regarded [30] as the tangent space at the identity to the manifold $\mathrm{Diff}_+ = \mathrm{Diff}_+(\mathbb{S}^1)$ of real-analytic orientation-preserving diffeomorphisms of $\mathbb{S}^1$.

As the tangent space to $\mathrm{Homeo}_+ \supseteq \mathrm{Diff}_+$ as in Section 3, $ppsl_2$ contains the Witt algebra; this inclusion is explicitly computed in Theorem A.1. Conversely for $A \in \mathrm{PSL}_2$, the Fourier expansion of the normalized wavelets $\bar{\vartheta}_A$ (given in Theorem A.2) or hyperfans ψ_A (given in Proposition A.4), describes how $ppsl_2$ lies inside the topological closure of the Witt algebra.

The quotient manifold $\mathrm{Diff}_+/M\ddot{o}b$, has tangent space at the identity given by the span of L_n, for $n^2 > 1$, that is, the span of $e, f, h \in sl_2$ is naturally identified with the span of L_{-1}, L_0, L_{+1}. As likewise follows from Section 3, the quotient $\mathrm{Homeo}_+^n \approx \mathrm{Homeo}_+/M\ddot{o}b$ can be identified with the space $\mathcal{T}ess^n$ of all normalized tesselations of $\mathbb{D}$, and the natural bundle $\widetilde{\mathcal{T}ess^n} \to \mathcal{T}ess^n$ over this space admits global affine lambda length coordinates.

There are the following several 2-forms defined on these spaces, which are compared in this section:

- The (pull-back from $\mathcal{T}ess^n$ of the) universal Weil-Petersson (WP) Kähler 2-form [25, 24] to $\widetilde{\mathcal{T}ess^n} \approx \widetilde{\mathrm{Homeo}_+^n}$ is given by

$$\omega = -2\sum d\mathrm{log}a \wedge d\mathrm{log}b \ + \ d\mathrm{log}b \wedge d\mathrm{log}c \ + \ d\mathrm{log}c \wedge d\mathrm{log}a,$$

where the sum is over the set of all triangles complementary to τ_* in $\mathbb{D}$ and the triangle has edges, here conflated with lambda lengths as usual, in the clockwise order a, b, c in the orientation on the underlying surface. This sum converges provided the homeomorphism is $C^{\frac{3}{2}+\epsilon}$ smooth on $\mathbb{S}^1$, cf. [32].

- The Kirillov-Kostant (KK) form [18, 32] is defined on $\mathrm{Diff}_+/M\ddot{o}b$ by

$$\kappa_a(L_m, L_n) = a(m^3 - m)\delta_{m,-n},$$

where $a \in \mathbb{C}$ and δ is the Kronecker delta function.

- The natural loop-algebra (LA) 2-cocycle [13] is defined by

$$\gamma(\psi^1, \psi^2) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{tr}\bigg(\psi^1(\theta) \cdot d\psi^2(\theta)\bigg),$$

where $\psi_1, \psi_2 \in sl_2^{\mathbb{S}^1}$ and tr denotes the trace. Our algebra $ppsl_2$ is a sub-algebra of a certain completion of $sl_2^{\mathbb{S}^1}$ on which this 2-cocycle still makes sense, and indeed integration by parts provides in this case the explicit expression

$$\gamma(\psi^1, \psi^2) = \frac{1}{2}\sum_{\theta \in \Pi(\psi^2)} \mathrm{tr}\bigg\{[\psi^1(\theta^+) + \psi^1(\theta^-)]\ [\psi^2(\theta^+) - \psi^2(\theta^-)]\bigg\}$$

for the 2-cocycle γ on $ppsl_2$, where $\Pi(\psi) \subset \hat{\mathbb{Q}} \subset \mathbb{S}^1$ denotes the set of breakpoints of ψ and $\theta^\pm$ denotes a point slightly $\frac{\text{beyond}}{\text{before}}$ the point

$\theta \in \Pi(\psi)$ in the counter-clockwise orientation on $\mathbb{S}^1$. One can verify directly that γ is a 2-cocycle by checking skew-symmetry and the 2-cocycle property, which follows from the 2-cocycle property for the first expression in this paragraph for γ on $sl_2^{\mathbb{S}^1}$.

In fact, these three 2-forms are pairwise identical up to overall constants, as we shall discuss. The constant for WP and LA will be computed here, and the constant for KK and WP was already calculated in Theorem 5.5 of [24], namely, we have

Theorem 6.1. *The WP Kähler 2-form* ω *and the KK form* κ_a *are related by*

$$\omega = \kappa_a, \text{ for } a = 2\pi i.$$

The computational proof in [24] using the basis of normalized arithmetic wavelets is involved and delicate depending upon the constraints on small Fourier modes described in Remark A.3.

Theorem 6.2. *The WP Kähler 2-form* ω *and the LA cocycle* γ *are related by*

$$\gamma = -4\omega.$$

Proof. Recall that $\Pi(\psi) \subset \hat{\mathbb{Q}} \subset \mathbb{S}^1$ denotes the (finite) set of breakpoints of $\psi \in ppsl_2$ and let $\bar{\vartheta}_A = \vartheta_A - X_A$, i.e., $\bar{\vartheta}_A$ is the normalized arithmetic wavelet and ϑ_A the unnormalized one. In particular, we have

$$\Pi(\bar{\vartheta}_A) = \Pi(\vartheta_A) = \Pi(\vartheta).A, \text{ for } A \in \mathrm{PSL}_2,$$

and moreover

$$\gamma(\bar{\vartheta}_A, \bar{\vartheta}_B) = \gamma(\bar{\vartheta}_A, \vartheta_B), \text{ for } A, B \in sl_2,$$

since X_B is both added and subtracted in a difference of $\bar{\vartheta}_B$-values in their contribution to γ. Thus

$$\gamma(\bar{\vartheta}_A, \bar{\vartheta}_B) = \gamma(\vartheta_A, \vartheta_B), \text{ for } A, B \in sl_2$$

as well by skew symmetry of γ.

We first claim that $\gamma(\bar\vartheta_A, \bar\vartheta_B) = 0$ if $\bar\vartheta_A$ takes a common value $Y_A \in sl_2$ at each point of $\Pi(\bar\vartheta_B)$. To see this, we compute

$$\begin{aligned}
\gamma(\bar\vartheta_A, \bar\vartheta_B) &= \gamma(\bar\vartheta_A, \vartheta_B) = \gamma(Y_A, \vartheta_B)\\
&= \sum_{\theta\in\Pi(\vartheta_B)} \operatorname{tr}\Big\{Y_A \cdot [\vartheta_B(\theta^+) - \vartheta_B(\theta^-)]\Big\}\\
&= \sum_{\theta\in\Pi(\vartheta)} \operatorname{tr}\Big\{Y_A \cdot [\vartheta_B(\theta^+.B) - \vartheta_B(\theta^-.B)]\Big\}\\
&= \operatorname{tr} \sum_{\theta\in\Pi(\vartheta)} Y_A\Big\{B^{-1}[\vartheta(\theta^+) - \vartheta(\theta^-)]B\Big\}\\
&= \operatorname{tr}\Big\{Y_A B^{-1}\Big(\sum_{\theta\in\Pi(\vartheta)} [\vartheta(\theta^+) - \vartheta(\theta^-)]\Big)B\Big\}\\
&= \operatorname{tr}\Big\{Y_A B^{-1}[4e + 2(f-e-h) - 4f + 2(h+f-e)]B\Big\}\\
&= \operatorname{tr}\, Y_A B^{-1} 0 B\\
&= 0.
\end{aligned}$$

It remains to consider $\gamma(\bar\vartheta_A, \bar\vartheta_{UA})$ and $\gamma(\bar\vartheta_A, \bar\vartheta_{TA})$, and we begin with the former. Let ξ_B, η_B denote the respective initial and terminal point of $e_B = e_I.B$, where e_I is the doe of τ_* as usual, so

$$\begin{aligned}
\Pi(\bar\vartheta_A) &= \{\xi_A, \eta_A, \eta_{U^{-1}A}, \eta_{UA}\},\\
\Pi(\bar\vartheta_{UA}) &= \{\xi_A, \eta_A, \eta_{UA}, \eta_{U^2A}\}.
\end{aligned}$$

Thus

$$\begin{aligned}
\gamma &= \gamma(\bar\vartheta_A, \bar\vartheta_{UA}) = \gamma(\vartheta_A, \vartheta_{UA})\\
&= \frac{1}{2} \sum_{\theta\in\Pi(\vartheta_{UA})} \operatorname{tr}\{[\vartheta_A(\theta^+) + \vartheta_A(\theta^-)][\vartheta_A(\theta^+) - \vartheta_A(\theta^-)]\},
\end{aligned}$$

so

$$2\gamma = \mathrm{tr}\left\{\begin{array}{l} \quad[\vartheta_A(\xi^+.A) + \vartheta_A(\xi^-.A)]\ [\vartheta_{UA}(\xi^+.A) - \vartheta_{UA}(\xi^-.A)] \\ +[\vartheta_A(\eta^+.A) + \vartheta_A(\eta^-.A)]\ [\vartheta_{UA}(\eta^+.A) - \vartheta_{UA}(\eta^-.A)] \\ +[\vartheta_A(\eta^+.UA) + \vartheta_A(\eta^-.UA)]\ [\vartheta_{UA}(\eta^+.UA) - \vartheta_{UA}(\eta^-.UA)] \\ +2[\vartheta_A(\eta^+.U^2A)]\ [\vartheta_{UA}(\eta^+.U^2A) - \vartheta_{UA}(\eta^-.U^2A)] \end{array}\right\}$$

$$= \mathrm{tr}\left\{\begin{array}{l} \quad[\vartheta(\xi^+) + \vartheta(\xi^-)]\ [(U)^{-1}\vartheta(\xi^+.U^{-1})U - (U)^{-1}\vartheta(\xi^-.U^{-1})U] \\ +[\vartheta(\eta^+) + \vartheta(\eta^-)]\ [(U)^{-1}\vartheta(\eta^+.U^{-1})U - (U)^{-1}\vartheta(\eta^-.U^{-1})U] \\ +[\vartheta(\eta^+.U) + A^{-1}\vartheta(\eta^-.U)]\ [(U)^{-1}\vartheta(\eta^+)U - (U)^{-1}\vartheta(\eta^-)U] \\ +2[\vartheta(\eta^+.U^2)]\ [(U)^{-1}\vartheta(\eta^+.U)U - (U)^{-1}\vartheta(\eta^-.U)U] \end{array}\right\}$$

$$= \mathrm{tr}\left\{\begin{array}{l} \quad\left[\begin{pmatrix}-1&0\\-2&1\end{pmatrix} + \begin{pmatrix}-1&0\\2&1\end{pmatrix}\right]\begin{pmatrix}1&0\\-1&1\end{pmatrix}\left[\begin{pmatrix}-1&0\\-2&1\end{pmatrix} - \begin{pmatrix}-1&0\\2&1\end{pmatrix}\right]\begin{pmatrix}1&0\\1&1\end{pmatrix} \\ +\left[\begin{pmatrix}1&2\\0&-1\end{pmatrix} + \begin{pmatrix}1&-2\\0&-1\end{pmatrix}\right]\begin{pmatrix}1&0\\-1&1\end{pmatrix}\left[\begin{pmatrix}1&-2\\0&-1\end{pmatrix} - \begin{pmatrix}-1&0\\-2&1\end{pmatrix}\right]\begin{pmatrix}1&0\\1&1\end{pmatrix} \\ +\left[\begin{pmatrix}-1&0\\2&1\end{pmatrix} + \begin{pmatrix}1&2\\0&-1\end{pmatrix}\right]\begin{pmatrix}1&0\\-1&1\end{pmatrix}\left[\begin{pmatrix}1&2\\0&-1\end{pmatrix} - \begin{pmatrix}1&-2\\0&-1\end{pmatrix}\right]\begin{pmatrix}1&0\\1&1\end{pmatrix} \\ +2\begin{pmatrix}-1&0\\2&1\end{pmatrix}\begin{pmatrix}1&0\\-1&1\end{pmatrix}\left[\begin{pmatrix}-1&0\\2&1\end{pmatrix} - \begin{pmatrix}1&2\\0&-1\end{pmatrix}\right]\begin{pmatrix}1&0\\1&1\end{pmatrix} \end{array}\right\}$$

$$\begin{aligned} &= \mathrm{tr}\left\{\begin{pmatrix}0&0\\-8&0\end{pmatrix} + \begin{pmatrix}0&-4\\-8&0\end{pmatrix} + \begin{pmatrix}-8&-8\\8&8\end{pmatrix} + 2\begin{pmatrix}4&2\\0&0\end{pmatrix}\right\} \\ &= 8, \end{aligned}$$

and accounting for the orientation and -2 in the expression above for the WP form yields the asserted constant $\gamma = -4\omega$.

Finally in the remaining case to compute $\gamma(\bar{\vartheta}_A, \bar{\vartheta}_{TA})$, replace $A \in \mathrm{PSL}_2$ by SUA, so that

$$\gamma(\bar{\vartheta}_{SUA}, \bar{\vartheta}_{TSUA}) = \gamma(\bar{\vartheta}_{UA}, \bar{\vartheta}_A) = +8$$

by skew symmetry since $\bar{\vartheta}_{SB} = \bar{\vartheta}_B$, for $B \in \mathrm{PSL}_2$, and

$$TSUA = TT^{-1}UT^{-1}UA = UT^{-1}UA = SA.$$

$\square$

7. Universal automorphic 1-form on $\mathrm{PPSL}_2(\mathbb{R})$

Given a decoration $\tilde{\tau}$ on a tesselation τ with framing framing $\mathcal{F}$, define

$$\xi_{\mathcal{F}} = \frac{1}{2} \sum_{A \in \mathrm{PSL}_2} \bar{\vartheta}_A \; d\log \lambda_{e_A} \in \Omega^1(\widetilde{\mathrm{Homeo}}{}^n_+, ppsl_2),$$

where $\bar{\vartheta}_A$ is the wavelet ϑ_A normalized with respect to $\mathcal{F}$, and λ_e denotes the lambda length of $e = e_A$. Alternatively, we may write $\xi_{\mathcal{F}} = \sum_{e \in \tau_*} \bar{\vartheta}_e \; d\log \lambda_e$ since $\vartheta_A = \vartheta_{SA}$ and $\lambda_e = \lambda_{e_A} = \lambda_{e_{SA}}$.

$\xi_{\mathcal{F}}$ is to be apprehended as a 1-form on the group $\widetilde{\mathrm{Homeo}}{}^n_+$ taking values in its Lie algebra $ppsl_2$. It evidently could be interpreted as the Maurer-Cartan form since the result of applying it to a multiplicative deformation of the lambda length, or in other words an additive deformation of the logarithm of the lambda length, is precisely the corresponding normalized vector field on $\mathbb{S}^1$ in its Lie algebra $ppsl_2$, as was discussed in Section 3.

More precisely, we have the diagram

$$\begin{tikzcd}
 & \widetilde{\mathrm{Homeo}}{}^n_+ \arrow[r, phantom, "\approx"] \arrow[d, "\pi"] & \widetilde{\mathrm{Homeo}}{}^n_+/M\ddot{o}b \arrow[d] \\
\mathrm{Homeo}_+ \arrow[r, shift left, "p"] & \mathrm{Homeo}^n_+ \arrow[l, shift left, "s_{\mathcal{F}}"] \arrow[r, phantom, "\approx"] & \mathrm{Homeo}_+/M\ddot{o}b,
\end{tikzcd}$$

where π is the forgetful map, p is the projection given by precomposition with $L^{\frac{0}{1},\frac{1}{0},\frac{1}{1}}_{f(\frac{0}{1}),f(\frac{1}{0}),f(\frac{1}{1})}$, and for any framing $\mathcal{F} = (u, v, w)$, the section $s_{\mathcal{F}}$ of p is given by precomposition with $L^{u,v,w}_{\frac{0}{1},\frac{1}{0},\frac{1}{1}}$. It is a tautology that if ω is the Maurer-Cartan form of Homeo_+, then $\xi_{\mathcal{F}} = \pi^* \circ s^*_{\mathcal{F}} \, \omega$.

Theorem 7.1. *For any fixed framing $\mathcal{F} \in C_3$, the Lie-algebra valued 1-form $\xi_{\mathcal{F}} \in \Omega^1(\widetilde{\mathrm{Homeo}}_+, ppsl_2)$ is invariant under flips.*

Proof. As a Lie algebra valued 1-form on a group, it is described by right-translating from the cotangent plane of the identity, and our computations will take place there.

Recall that the Farey tesselation τ_* admits a canonical decoration $\tilde{\tau}_*$ determined by the property that all the lambda lengths are constant equal to unity, and that this is the identity element of the group $\widetilde{\mathrm{Homeo}}{}^n_+$ as discussed in Section 3.

To establish notation, refer to Figure 1 and consider the four generation-zero and -one Farey points $\frac{1}{0}, -\frac{1}{1}, \frac{0}{1}, \frac{1}{1}$ decomposing $\mathbb{S}^1$ into four circular intervals which are conflated with the respective quadrants I,II,III and

IV containing them. The convex hull of these points is an ideal quadrilateral with oriented frontier edges $a = e_{ST}, b = e_{SU}, c = e_{U^{-1}}$ and $d = e_{T^{-1}}$ in this counter-clockwise order starting from $\frac{1}{0}$. The doe e_I will be denoted simply $e = e_I$, and the edge arising from its flip, with respective initial and terminal points $\frac{1}{1}$ and $-\frac{1}{1}$, will be denoted f.

The second-generation $-\frac{2}{1}, -\frac{1}{2}, \frac{1}{2}, \frac{2}{1}$ Farey points further decompose the circular intervals $\mathrm{I} = \mathrm{I}_- \cup \mathrm{I}_+, \ldots, \mathrm{IV} = \mathrm{IV}_- \cup \mathrm{IV}_+$ occurring in this counter-clockwise order starting from $\frac{1}{0}$. The Farey points of generation at most two together span an ideal octagon with oriented frontier edges e_{ST^2}, e_{SUT}, e_{TSU}, e_{SU^2}, $e_{U^{-2}}$, $e_{T^{-1}U^{-1}}$, $e_{U^{-1}T^{-1}}$, and $e_{T^{-2}}$ in this counter-clockwise order starting from $\frac{1}{0}$, as indicated in Figure 2.

The hyperbolic transformations which are primitive in PSL_2 along the respective axes e and f are given by the exponentials of $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and along each of the other edges e_A above by the exponential of $X_A \in sl_2$ with

$$X_{ST} = \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right),\ X_{SU} = \left(\begin{smallmatrix} 1 & 0 \\ -2 & -1 \end{smallmatrix}\right),\ X_{U^{-1}} = \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right),\ X_{T^{-1}} = \left(\begin{smallmatrix} -1 & 2 \\ 0 & 1 \end{smallmatrix}\right),$$

and

$$\begin{aligned} X_{ST^2} &= \left(\begin{smallmatrix} 1 & 4 \\ 0 & -1 \end{smallmatrix}\right),\ X_{SUT} = \left(\begin{smallmatrix} 3 & 4 \\ -2 & -3 \end{smallmatrix}\right), \\ X_{TSU} &= \left(\begin{smallmatrix} 3 & 2 \\ -4 & -3 \end{smallmatrix}\right),\ X_{SU^2} = \left(\begin{smallmatrix} 1 & 0 \\ -4 & -1 \end{smallmatrix}\right), \\ X_{U^{-2}} &= \left(\begin{smallmatrix} -1 & 0 \\ -4 & 1 \end{smallmatrix}\right),\ X_{T^{-1}U^{-1}} = \left(\begin{smallmatrix} -3 & 2 \\ -4 & 3 \end{smallmatrix}\right), \\ X_{U^{-1}T^{-1}} &= \left(\begin{smallmatrix} -3 & 4 \\ -2 & 3 \end{smallmatrix}\right),\ X_{U^{-1}T^{-1}} = \left(\begin{smallmatrix} -1 & 4 \\ 0 & 1 \end{smallmatrix}\right). \end{aligned}$$

The proof is a direct computation and is omitted.

These are the logarithms of the component earthquakes for the wavelets ϑ_x, for $x \in \{a, b, c, d, e, f\}$, so it is not difficult to combine them four at a time and normalize with respect to the the standard framing $\mathcal{F} = (\frac{0}{1}, \frac{1}{0}, \frac{1}{1})$ on τ_*. One finds

$$\bar{\vartheta}_e = \vartheta_e = \begin{cases} \left(\begin{smallmatrix} -1 & -2 \\ 0 & 1 \end{smallmatrix}\right), & \text{on I;} \\ \left(\begin{smallmatrix} 1 & 0 \\ -2 & -1 \end{smallmatrix}\right), & \text{on II;} \\ \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right), & \text{on III;} \\ \left(\begin{smallmatrix} -1 & 2 \\ 0 & 1 \end{smallmatrix}\right), & \text{on IV,} \end{cases}$$

$$\vartheta_a = \begin{cases} \left(\begin{smallmatrix} -1 & -4 \\ 0 & 1 \end{smallmatrix}\right), & \text{on I}_-; \\ \left(\begin{smallmatrix} 3 & 4 \\ -2 & -3 \end{smallmatrix}\right), & \text{on I}_+; \\ \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right), & \text{on II;} \\ \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), & \text{on III and IV,} \end{cases} \qquad \text{so } \bar{\vartheta}_a = \begin{cases} \left(\begin{smallmatrix} 0 & -4 \\ 0 & 0 \end{smallmatrix}\right), & \text{on I}_-; \\ \left(\begin{smallmatrix} 4 & 4 \\ -2 & -4 \end{smallmatrix}\right), & \text{on I}_+; \\ \left(\begin{smallmatrix} 0 & 0 \\ 2 & 0 \end{smallmatrix}\right), & \text{on II;} \\ \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right), & \text{on III and IV,} \end{cases}$$

$$\vartheta_b = \begin{cases} \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right), & \text{on I;} \\ \left(\begin{smallmatrix} -3 & -2 \\ 4 & 3 \end{smallmatrix}\right), & \text{on II}_-; \\ \left(\begin{smallmatrix} 1 & 0 \\ -4 & -1 \end{smallmatrix}\right), & \text{on II}_+; \\ \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), & \text{on III and IV,} \end{cases} \quad \text{so } \bar{\vartheta}_b = \begin{cases} \left(\begin{smallmatrix} 0 & 2 \\ 0 & 0 \end{smallmatrix}\right), & \text{on I;} \\ \left(\begin{smallmatrix} -4 & -2 \\ 4 & 4 \end{smallmatrix}\right), & \text{on II}_-; \\ \left(\begin{smallmatrix} 0 & 0 \\ -4 & 0 \end{smallmatrix}\right), & \text{on II}_+; \\ \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right), & \text{on III and IV,} \end{cases}$$

$$\bar{\vartheta}_c = \vartheta_c = \begin{cases} \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), & \text{on I and II;} \\ \left(\begin{smallmatrix} 1 & 0 \\ 4 & -1 \end{smallmatrix}\right), & \text{on III}_-; \\ \left(\begin{smallmatrix} -3 & 2 \\ -4 & 3 \end{smallmatrix}\right), & \text{on III}_+; \\ \left(\begin{smallmatrix} 1 & -2 \\ 0 & -1 \end{smallmatrix}\right), & \text{on IV,} \end{cases} \quad \text{and } \bar{\vartheta}_d = \vartheta_d = \begin{cases} \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), & \text{on I and II;} \\ \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right), & \text{on III;} \\ \left(\begin{smallmatrix} 3 & -4 \\ 2 & -3 \end{smallmatrix}\right), & \text{on IV}_-; \\ \left(\begin{smallmatrix} -1 & 4 \\ 0 & 1 \end{smallmatrix}\right), & \text{on IV}_+ \end{cases}$$

As illustrated in Figure 5, letting $\tilde{x} = d\log\lambda_x$, for $x \in \{a, b, c, d, e, f\}$, and putting all of this together, we find that

$$\begin{aligned} &\sum_{x\in\{a,b,c,d,e\}} \bar{\vartheta}_x \; d\log\lambda_x \\ &= \begin{cases} \left(\begin{smallmatrix} 0 & -4 \\ 0 & 0 \end{smallmatrix}\right)\tilde{a} + \left(\begin{smallmatrix} 0 & 2 \\ 0 & 0 \end{smallmatrix}\right)\tilde{b} + \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} -1 & -2 \\ 0 & 1 \end{smallmatrix}\right)\tilde{e}, & \text{on I}_-; \\ \left(\begin{smallmatrix} 4 & 4 \\ -2 & -4 \end{smallmatrix}\right)\tilde{a} + \left(\begin{smallmatrix} 0 & 2 \\ 0 & 0 \end{smallmatrix}\right)\tilde{b} + \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} -1 & -2 \\ 0 & 1 \end{smallmatrix}\right)\tilde{e}, & \text{on I}_+; \\ \left(\begin{smallmatrix} 0 & 0 \\ 2 & 0 \end{smallmatrix}\right)\tilde{a} + \left(\begin{smallmatrix} -4 & -2 \\ 4 & 4 \end{smallmatrix}\right)\tilde{b} + \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} 1 & 0 \\ -2 & -1 \end{smallmatrix}\right)\tilde{e}, & \text{on II}_-; \\ \left(\begin{smallmatrix} 0 & 0 \\ 2 & 0 \end{smallmatrix}\right)\tilde{a} + \left(\begin{smallmatrix} 0 & 0 \\ -4 & 0 \end{smallmatrix}\right)\tilde{b} + \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} 1 & 0 \\ -2 & -1 \end{smallmatrix}\right)\tilde{e}, & \text{on II}_+; \\ \left(\begin{smallmatrix} 1 & 0 \\ 4 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right)\tilde{e}, & \text{on III}_-; \\ \left(\begin{smallmatrix} -3 & 2 \\ -4 & 3 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right)\tilde{e}, & \text{on III}_+; \\ \left(\begin{smallmatrix} 1 & -2 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} 3 & -4 \\ 2 & -3 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} -1 & 2 \\ 0 & 1 \end{smallmatrix}\right)\tilde{e}, & \text{on IV}_-; \\ \left(\begin{smallmatrix} 1 & -2 \\ 0 & -1 \end{smallmatrix}\right)\tilde{c} + \left(\begin{smallmatrix} -1 & 4 \\ 0 & 1 \end{smallmatrix}\right)\tilde{d} + \left(\begin{smallmatrix} -1 & 2 \\ 0 & 1 \end{smallmatrix}\right)\tilde{e}, & \text{on IV}_+. \end{cases} \end{aligned}$$

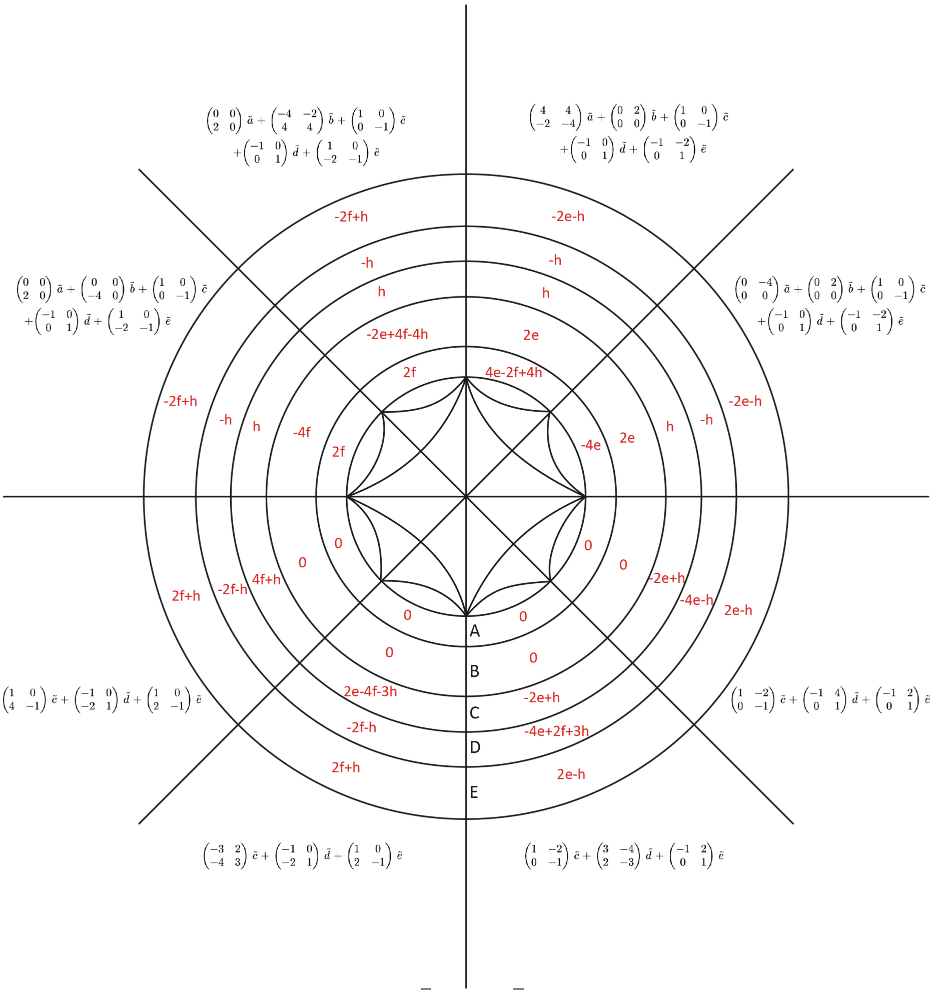

$\begin{pmatrix}0&0\\2&0\end{pmatrix}\tilde{a}+\begin{pmatrix}-4&-2\\4&4\end{pmatrix}\tilde{b}+\begin{pmatrix}1&0\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\0&1\end{pmatrix}\tilde{d}+\begin{pmatrix}1&0\\-2&-1\end{pmatrix}\tilde{e}$
$\begin{pmatrix}4&4\\-2&-4\end{pmatrix}\tilde{a}+\begin{pmatrix}0&2\\0&0\end{pmatrix}\tilde{b}+\begin{pmatrix}1&0\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\0&1\end{pmatrix}\tilde{d}+\begin{pmatrix}-1&-2\\0&1\end{pmatrix}\tilde{e}$
$\begin{pmatrix}0&0\\2&0\end{pmatrix}\tilde{a}+\begin{pmatrix}0&0\\-4&0\end{pmatrix}\tilde{b}+\begin{pmatrix}1&0\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\0&1\end{pmatrix}\tilde{d}+\begin{pmatrix}1&0\\-2&-1\end{pmatrix}\tilde{e}$
$\begin{pmatrix}0&-4\\0&0\end{pmatrix}\tilde{a}+\begin{pmatrix}0&2\\0&0\end{pmatrix}\tilde{b}+\begin{pmatrix}1&0\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\0&1\end{pmatrix}\tilde{d}+\begin{pmatrix}-1&-2\\0&1\end{pmatrix}\tilde{e}$
-2f+h
-2e-h
-h
-h
h
h
-2e+4f-4h
2e
2f
4e-2f+4h
-2f+h
-h
h
-4f
2f
-4e
2e
h
-h
-2e-h
0
0
0
0
-2e+h
-4e-h
2e-h
2f+h
-2f-h
4f+h
0
0
A
0
0
B
$\begin{pmatrix}1&0\\4&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\-2&1\end{pmatrix}\tilde{d}+\begin{pmatrix}1&0\\2&-1\end{pmatrix}\tilde{e}$
2e-4f-3h
-2e+h
$\begin{pmatrix}1&-2\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&4\\0&1\end{pmatrix}\tilde{d}+\begin{pmatrix}-1&2\\0&1\end{pmatrix}\tilde{e}$
C
-2f-h
-4e+2f+3h
D
2f+h
2e-h
E
$\begin{pmatrix}-3&2\\-4&3\end{pmatrix}\tilde{c}+\begin{pmatrix}-1&0\\-2&1\end{pmatrix}\tilde{d}+\begin{pmatrix}1&0\\2&-1\end{pmatrix}\tilde{e}$
$\begin{pmatrix}1&-2\\0&-1\end{pmatrix}\tilde{c}+\begin{pmatrix}3&-4\\2&-3\end{pmatrix}\tilde{d}+\begin{pmatrix}-1&2\\0&1\end{pmatrix}\tilde{e}$

Figure 5. Depiction of $\bar{\vartheta}_a, \ldots, \bar{\vartheta}_e$ in the respective rings $A, \ldots, E$ and $\sum_{x\in\{a,b,c,d,e\}} \bar{\vartheta}_x \; d\log\lambda_x$ on the outside.

Let τ' denote the tesselation that arises from τ_* by a flip along the edge $e \in \tau_*$, let ϑ'_x denote the vectorfield on $\mathbb{S}^1$ corresponding in the natural way to the edge $x \in \{a, b, c, d, f\} \subseteq \tau'$, and let $\bar{\vartheta}'_x$ denote the normalization of ϑ'_x relative to the same framing $\mathcal{F}$. Thus, we have

$$\bar{\vartheta}'_f = \vartheta'_f = \bar{\vartheta}_e,$$

and a further consequence of the primitives for logarithms of hyperbolic transformations computed before is that

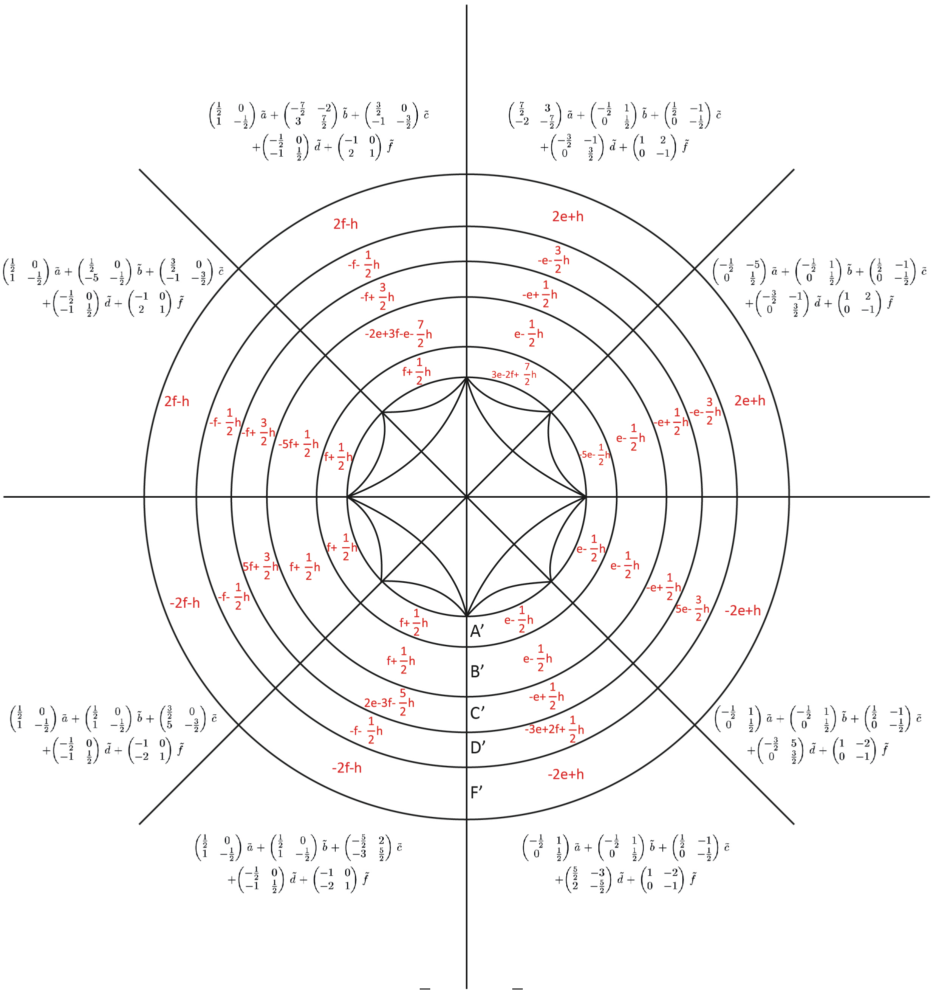

$\begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} -\frac{7}{2} & -2 \\ 3 & \frac{7}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{3}{2} & 0 \\ -1 & -\frac{3}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} \frac{7}{2} & 3 \\ -2 & -\frac{7}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{3}{2} & -1 \\ 0 & \frac{3}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} \frac{1}{2} & 0 \\ -5 & -\frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{3}{2} & 0 \\ -1 & -\frac{3}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} -\frac{1}{2} & -5 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{3}{2} & -1 \\ 0 & \frac{3}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \tilde{f}$
2f-h
2e+h
$-f-\frac{1}{2}h$
$-e-\frac{3}{2}h$
$-f+\frac{3}{2}h$
$-e+\frac{1}{2}h$
$-2e+3f-e-\frac{7}{2}h$
$e-\frac{1}{2}h$
$f+\frac{1}{2}h$
$3e-2f+\frac{7}{2}h$
2f-h
$-f-\frac{1}{2}h$
$-f+\frac{3}{2}h$
$-5f+\frac{1}{2}h$
$f+\frac{1}{2}h$
$-5e-\frac{1}{2}h$
$e-\frac{1}{2}h$
$-e+\frac{1}{2}h$
$-e-\frac{3}{2}h$
2e+h
-2f-h
$-f-\frac{1}{2}h$
$5f+\frac{3}{2}h$
$f+\frac{1}{2}h$
$f+\frac{1}{2}h$
$e-\frac{1}{2}h$
$e-\frac{1}{2}h$
$-e+\frac{1}{2}h$
$5e-\frac{3}{2}h$
-2e+h
$f+\frac{1}{2}h$
$e-\frac{1}{2}h$
A′
$f+\frac{1}{2}h$
$e-\frac{1}{2}h$
B′
$2e-3f-\frac{5}{2}h$
$-e+\frac{1}{2}h$
C′
$-f-\frac{1}{2}h$
$-3e+2f+\frac{1}{2}h$
D′
-2f-h
-2e+h
F′
$\begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{3}{2} & 0 \\ 5 & -\frac{3}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{3}{2} & 5 \\ 0 & \frac{3}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} -\frac{5}{2} & 2 \\ -3 & \frac{5}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} \tilde{f}$
$\begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{a} + \begin{pmatrix} -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \tilde{b} + \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} \end{pmatrix} \tilde{c} + \begin{pmatrix} \frac{5}{2} & -3 \\ 2 & -\frac{5}{2} \end{pmatrix} \tilde{d} + \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \tilde{f}$

Figure 6. Depiction of $\bar{\vartheta}'_a, \ldots, \bar{\vartheta}'_f$ in the respective rings $A', \ldots D', F'$ and $\sum_{x\in\{a,b,c,d,f\}} \bar{\vartheta}'_x \; d\log\lambda_x$ on the outside.

$$\bar{\vartheta}'_a = \vartheta'_a + \begin{pmatrix} \frac{1}{2} & -1 \\ 0 & -\frac{1}{2} \end{pmatrix}, \text{ where } \vartheta'_a = \begin{cases} \left(\begin{smallmatrix} -1 & -4 \\ 0 & 1 \end{smallmatrix}\right), & \text{on I}_-; \\ \left(\begin{smallmatrix} 3 & 4 \\ -2 & -3 \end{smallmatrix}\right), & \text{on I}_+; \\ \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), & \text{on II and III;} \\ \left(\begin{smallmatrix} -1 & 2 \\ 0 & 1 \end{smallmatrix}\right), & \text{on IV,} \end{cases}$$

$$\bar{\vartheta}'_b = \vartheta'_b + \begin{pmatrix} -\frac{1}{2} & 0 \\ -1 & \frac{1}{2} \end{pmatrix}, \text{ where } \vartheta'_b = \begin{cases} \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), & \text{on I and IV;} \\ \left(\begin{smallmatrix} -3 & -2 \\ 4 & 3 \end{smallmatrix}\right), & \text{on II}_-; \\ \left(\begin{smallmatrix} 1 & 0 \\ -4 & -1 \end{smallmatrix}\right), & \text{on II}_+; \\ \left(\begin{smallmatrix} 1 & 0 \\ 2 & -1 \end{smallmatrix}\right), & \text{on III,} \end{cases}$$

$$\bar\vartheta'_c = \vartheta'_c + \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}, \text{ where } \vartheta'_c = \begin{cases} \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right), & \text{on I and IV;} \\ \left(\begin{smallmatrix} 1 & 0 \\ -2 & -1 \end{smallmatrix}\right), & \text{on II;} \\ \left(\begin{smallmatrix} 1 & 0 \\ 4 & -1 \end{smallmatrix}\right), & \text{on III}_-; \\ \left(\begin{smallmatrix} -3 & 2 \\ -4 & -1 \end{smallmatrix}\right), & \text{on III}_+, \end{cases}$$

$$\bar\vartheta'_d = \vartheta'_d + \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}, \text{ where } \vartheta'_d = \begin{cases} \left(\begin{smallmatrix} -1 & -2 \\ 0 & 1 \end{smallmatrix}\right), & \text{on I,;} \\ \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right), & \text{on II and III;} \\ \left(\begin{smallmatrix} -3 & 4 \\ -2 & 3 \end{smallmatrix}\right), & \text{on IV}_-; \\ \left(\begin{smallmatrix} -1 & 4 \\ 0 & 1 \end{smallmatrix}\right), & \text{on IV}_+. \end{cases}$$

Again putting all this together as illustrated in Figure 6, we find

$$\sum_{x\in\{a,b,c,d,f\}} \bar\vartheta'_x \; d\log\lambda_x$$
$$= \begin{cases} \begin{pmatrix} -\frac12 & -5 \\ 0 & \frac12 \end{pmatrix}\tilde a + \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac12 & -1 \\ 0 & -\frac12 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac32 & -1 \\ 0 & \frac32 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right)\tilde f, & \text{on I}_-; \\ \begin{pmatrix} \frac72 & 3 \\ -2 & -\frac12 \end{pmatrix}\tilde a + \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac12 & -1 \\ 0 & -\frac12 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac32 & -1 \\ 0 & \frac32 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right)\tilde f, & \text{on I}_+; \\ \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde a + \begin{pmatrix} -\frac72 & -2 \\ 3 & \frac72 \end{pmatrix}\tilde b + \begin{pmatrix} \frac32 & 0 \\ -1 & -\frac32 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac12 & 0 \\ -1 & \frac12 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} -1 & 0 \\ 2 & 1 \end{smallmatrix}\right)\tilde f, & \text{on II}_-; \\ \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde a + \begin{pmatrix} \frac12 & 0 \\ -5 & -\frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac32 & 0 \\ -1 & -\frac32 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac12 & 0 \\ -1 & \frac12 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} -1 & 0 \\ 2 & 1 \end{smallmatrix}\right)\tilde f, & \text{on II}_+; \\ \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde a + \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac32 & 0 \\ 5 & -\frac32 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac12 & 0 \\ -1 & \frac12 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right)\tilde f, & \text{on III}_-; \\ \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde a + \begin{pmatrix} \frac12 & 0 \\ 1 & -\frac12 \end{pmatrix}\tilde b + \begin{pmatrix} -\frac52 & 2 \\ -3 & \frac52 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac12 & 0 \\ -1 & \frac12 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} -1 & 0 \\ -2 & -1 \end{smallmatrix}\right)\tilde f, & \text{on III}_+; \\ \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde a + \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac12 & -1 \\ 0 & -\frac12 \end{pmatrix}\tilde c + \begin{pmatrix} \frac52 & -3 \\ 2 & -\frac52 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} 1 & -2 \\ 0 & -1 \end{smallmatrix}\right)\tilde f, & \text{on IV}_-; \\ \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde a + \begin{pmatrix} -\frac12 & 1 \\ 0 & \frac12 \end{pmatrix}\tilde b + \begin{pmatrix} \frac12 & -1 \\ 0 & -\frac12 \end{pmatrix}\tilde c + \begin{pmatrix} -\frac32 & 5 \\ 0 & \frac32 \end{pmatrix}\tilde d + \left(\begin{smallmatrix} 1 & -2 \\ 0 & -1 \end{smallmatrix}\right)\tilde f, & \text{on IV}_+; \end{cases}$$

Using the fact that

$$\tilde e + \tilde f = \frac{1}{ac+bd}(\tilde a + \tilde b + \tilde c + \tilde d) = \frac12(\tilde a + \tilde b + \tilde c + \tilde d)$$

by the Ptolemy relation $ef = ac + bd$, since $a = b = c = d = 1$ for the identity element $\tilde\tau_*$, a computation with rather miraculous cancellations as illustrated in Figure 7 confirms that

$$\sum_{x\in\{a,b,c,d,e\}} \bar\vartheta_x \; d\log\lambda_x \;=\; \sum_{x\in\{a,b,c,d,e\}} \bar\vartheta'_x \; d\log\lambda_x,$$

as required to check invariance under the flip along the doe for the standard framing $\mathcal{F} = (\frac{0}{1}, \frac{1}{0}, \frac{1}{1})$.

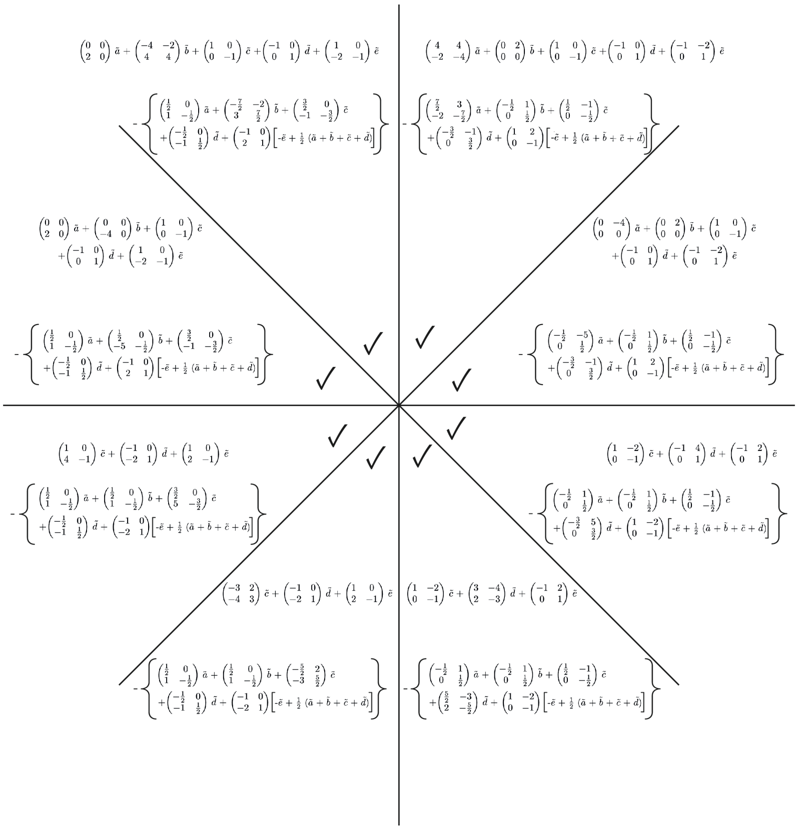


Figure 7. Difference between expressions in Figures 5 and 6 vanishes for all lambda lengths equal to unity, i.e., at the universal punctured arithmetic surface.

This establishes the invariance for a flip on the doe with the associated framing. In the general case first of all, the 1-form is invariant under push-forward by the Möbius group since lambda lengths are $M\ddot{o}b$-invariant, so we may assume that the framing is one that is associated to a doe. Likewise without loss of generality, we may assume that the flip is performed on the edge with endpoints $\frac{0}{1}, \frac{1}{0} \in \mathbb{S}^1$, with the doe however potentially located elsewhere.

Thus, the asserted invariance must be checked for the various possibilities of relative position of doe and the edge with endpoints $\frac{0}{1}, \frac{1}{0}$ to be flipped. In the presented computation, these two edges coincide, and there are four further cases to check depending upon which quadrant in $\mathbb{C} \supseteq \mathbb{D}$ contains the doe, thereby determining the normalization. These four similar but simpler calculations, each corresponding to a different normalization of the common unnormalized formulae already presented in the proof, are left as an exercise for the reader. $\square$

Consider an arbitrary framing $\mathcal{F}$ and let

$$\eta_\mathcal{F} = s^*_\mathcal{F}\, \omega \in \Omega^1(\mathrm{Homeo}_+/M\ddot{o}b),$$

which by general principles should satisfy

$$\underline{\text{Maurer} - \text{Cartan equation}}:\ 0 = d\eta_\mathcal{F} + \frac{1}{2}[\eta_\mathcal{F}, \eta_\mathcal{F}],\ \text{for any framing } \mathcal{F};$$

$$\underline{\text{Compatibility Equation}}: \eta_\mathcal{G} = \mathrm{Ad}(h^{-1}_{\mathcal{F}\mathcal{G}})\, \eta_\mathcal{F}\ +\ h^*_{\mathcal{F}\mathcal{G}}\, \omega_{M\ddot{o}b}, \text{for framings}$$

$\mathcal{F}, G,$ where $h_{\mathcal{F}\mathcal{G}} = s_\mathcal{G} \circ s^{-1}_\mathcal{F}$, Ad denotes the adjoint and $\omega_{M\ddot{o}b}$ is the Maurer-Cartan form on $M\ddot{o}b$.

Note that the Compatibility Equation together with the result for the flip on the doe with its associated framing would give an alternative proof of Theorem 7.1. This putative proof is a bit of a swindle because *there is no Maurer-Cartan form on the topological group* Homeo$_+$ except for the version here, and hence no a priori Compatibility Equation.

Remark 7.2. The expression for $\xi_\mathcal{F}$ as a Poincaré series in framed holographic coordinates is easily derived from Remark 2.8.

As we have noted, the form $\xi_\mathcal{F}$ is morally the Maurer-Cartan form in the sense that its value on the tangent vector $\frac{\partial}{\partial \ln \lambda_A}$ at the identity of $\mathrm{PPSL}_2(\mathbb{R})$ is the corresponding normalized element $\bar{\vartheta}_A$ of the Lie algebra $ppsl_2$. However, $\xi_\mathcal{F}$ has been shown here to be invariant under the lattice $\mathrm{PPSL}_2(\mathbb{Z})$ rather than under the full group $\mathrm{PPSL}_2(\mathbb{R})$, whose invariance we may conjecture. On the other hand, one might therefore more squarely regard $\xi_\mathcal{F}$ solely as an automorphic form, and this is the viewpoint of the next section.

8. Future perspectives

We conclude with a heuristic discussion of constructing the automorphic representation for our universal triad corresponding to the automorphic 1-form obtained in Section 7. To this end, we combine the original constructions of [34] and of Vaughan Jones [16, 17], also described in [22].

We recall from Appendix B that we have constructed an indecomposable automorphic representation V of $\mathrm{PSL}_2(\mathbb{R})$ by lifting the weight 2 Eisenstein series E_2. Viewed as the Lie algebra $psl_2 = psl_2(\mathbb{R})$ representation, it can be identified with the Laurent polynomial ring

$$V \approx \mathcal{O}(\mathbb{C}^\times) = \mathbb{C}[z^{\pm 1}].$$

with the psl_2 action given by

$$e = \frac{d}{dz},\ h = 2z\frac{d}{dz},\ f = z^2\frac{d}{dz}.$$

The space V has a natural multiplication map

$$V \to V \otimes V$$

which commutes with the action of psl_2. One can also identify the restricted dual space with Laurent 1-forms

$$V' \approx \Omega^1(\mathbb{C}^\times) = \mathbb{C}[z^{\pm 1}]dz,$$

with the pairing of V and V' given by the residue of the product at 0. This gives the dual map

$$V' \to V' \otimes V', \tag{8.1}$$

which can be used for the inductive limit construction as in [16, 17, 22].

Namely for any polygon P_γ with frontier γ, we assign a tensor product

$$H_\gamma = \bigotimes_{j=1}^{n} V',$$

where n is the number of edges in γ. There is a partial ordering on the set of polygons given by inclusion $P_\gamma \subseteq P_{\gamma'}$, which induces the partial ordering $\gamma \leqslant \gamma'$ on frontiers. Clearly the Lie subalgebra

$$ppsl_2(\gamma) \subseteq ppsl_2$$

of piecewise psl_2 maps with breakpoints at the vertices of γ acts on H_γ in the natural way, where psl_2 corresponding to the j^{th} edge acts on the j^{th} factor of H_γ. Then the map (8.1) induces a family of intertwining operators

$$T_\gamma^{\gamma'} : H_\gamma \to H_{\gamma'},$$

so that $T_\gamma^\gamma = 1$ and the consistency condition $T_{\gamma''}^{\gamma} = T_{\gamma''}^{\gamma'} T_{\gamma'}^{\gamma}$ holds whenever $\gamma \leqslant \gamma' \leqslant \gamma''$. The direct limit

$$H = \varinjlim H_\gamma \tag{8.2}$$

yields a representation of $ppsl_2$. It can be made unitary by means of a 1-cocycle determined by the one-dimensional subrepresentation $V_0 \subseteq V$ as in [34] or alternatively using a modification of the construction above based upon the approximation of the representation V' of psl_2 by the representations of the complementary series $V(\lambda)$, for $0 < \lambda < 2$ with $\lambda \neq 1$, so that $V(\lambda) \approx V(2-\lambda)$ with $\lim\limits_{\lambda \to 2} V(\lambda) = V$ and $\lim\limits_{\lambda \to 0} V(\lambda) = V'$; see [34] for details.

Remark 8.1. The partial ordering above on ideal polygons occurring as fundamental domains for punctured surfaces leads one to consider the so-called punctured solenoid $\mathcal{H}$, which is an initial object for the category of punctured surfaces with morphisms given by finite covers branched only over the punctures. The unpunctured version was introduced and studied by Dennis Sullivan in [31]. The decorated Teichmüller space of $\mathcal{H}$ was studied in [28] and parametrized by certain coordinates on the Farey tesselation τ_*, namely, lambda length *functions*, one such continuous and $\mathrm{PSL}_2(\mathbb{Z})$-equivariant function from the profinite completion of $\mathrm{PSL}_2(\mathbb{Z})$ to $\mathbb{R}_{>0}$ for each edge of τ_*. Flip-like generators for the mapping class group of $\mathcal{H}$ are derived in [28], and a complete set of relations in these generators is finally given in [2].

To implement our program, we want to realize H, or heuristically $(V')^{\otimes\infty}$, as an automorphic representation on $\mathrm{PPSL}_2(\mathbb{Z})\backslash\mathrm{PPSL}_2(\mathbb{R})$ of $ppsl_2$. To this end, we notice first that the space of harmonic functions on $\mathrm{PSL}_2(\mathbb{R})$, cf. (B.3), can be identified with the space of functions on the boundary of $\mathrm{PSL}_2(\mathbb{R})$, which is dual to the space of functions of the holographic coordinates $\{(s, \delta)\}$ introduced in Section 2. Furthermore for both spaces, we can also identify the corresponding $\mathrm{PSL}_2(\mathbb{R})$ actions, and therefore, we can realize V' by imposing modular invariance on the space of functions depending on a single pair (s, δ) of holographic coordinates. This realization clearly extends to finite tensor products $(V')^{\otimes n}$ associated to n-gons P_γ. We conjecture that the inductive limit H also admits a realization by the automorphic functions in holographic coordinates, with $\mathrm{PPSL}_2(\mathbb{Z})$ playing the role of the modular group $\mathrm{PSL}_2(\mathbb{Z})$.

On the other hand, the analogy between $\mathrm{PPSL}_2(\mathbb{Z})$ and the mapping class group suggests an extension of the former by an infinite symmetric group, or more generally by an infinite braid group, which is still a discrete subgroup of $\mathrm{Homeo}_+(\mathbb{S}^1)$, cf. [14] and the references therein. The reduction of H by the additional symmetry yields the space Λ, or heuristically $S^\infty V'$, which is naturally identified with the free bsosonic field, one of the simplest CFT_2.

As discussed in Appendix B, the theory of automorphic forms allows the enlargement of various spaces by relaxing the automorphic property to smaller discrete subgroups; for example, the restriction to invariance by the commutant $\mathrm{PSL}_2(\mathbb{Z})' \subseteq \mathrm{PSL}_2(\mathbb{Z})$ enlarges V sixfold additively. Similarly, we expect that replacing $\mathrm{PPSL}_2(\mathbb{Z})$ by $\mathrm{P}(\mathrm{PSL}_2(\mathbb{Z}))'$ will correspondingly yield a multiplicative increase of Λ and thereby allow the construction of a CFT_2 of size comparable to the Monster CFT_2.

Finally we come to the question of how we can effectively capture the Monster using the new universal automorphic theory, and in particular how we can derive the Monstrous Moonshine. At this moment, we do not know the complete answer, but we can begin on the opposite side, which will help to clarify the problem and capture the beast. Namely, we consider all 194 conjugacy classes of the Monster and the Thompson series of the representations of these classes in the the Monster representation of [12]. Then we know from [1] that they are the canonical Hauptmoduln $J_\Gamma(z)$ for some genus-zero discrete subgroups Γ of $\mathrm{PSL}_2(\mathbb{R})$ of Moonshine type, i.e., groups commensurable with $\mathrm{PSL}_2(\mathbb{Z})$ which contain $\Gamma_\infty = \{\pm\left(\begin{smallmatrix}1 & n\\ 0 & 1\end{smallmatrix}\right)\}$. Then one has

$$J_\Gamma(z) = q^{-1} + \sum_{n=1}^{\infty} c_\Gamma(n) q^n,$$

for some integral coefficients $c_\Gamma(n)$. One can also construct the corresponding Eisenstein series using a logarithmic derivative of $J_\Gamma(z)$, or as usual an analytic continuation to $s = 0$ of the series

$$E_2^\Gamma(z,s) = \sum_{(c,d)\in\Gamma/\Gamma_\infty} \frac{(\mathrm{Im}\ z)^s}{(cz+d)^2|cz+d|^{2s}}.$$

As in the classical example $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, the resulting Eisenstein series after multiplication by dz is Γ-invariant though not holomorphic. Then we can repeat the lift of a Γ-invariant 1-form recalled in Appendix B to the automorphic function

$$f = E_2^\Gamma \rightsquigarrow \phi_f(g) = (ci+d)^{-2} f(g\cdot i),$$

where $g = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{R})$, and we again obtain an indecomposable representation of $\mathrm{PSL}_2(\mathbb{R})$ that we shall denote V^Γ. In fact, we just get another model of the same indecomposable representation of $\mathrm{PSL}_2(\mathbb{R})$. We can also expect a similar construction of the space of universal automorphic forms for $\mathrm{P}\Gamma$ denoted Λ^Γ. It is natural to conjecture that we find a twisted conformal field theory associated to Λ^Γ and the Monster conjugacy class corresponding to Γ. The comparison of Λ and Λ^Γ should yield the Monster element in this conjugacy class, and they together will allow realization of the Monster group via automorphisms of Λ. At this point, the trap will close upon the Monster, and the Moonshine will follow.

Appendix A. Harmonic analysis

To begin, we express the generators of the Witt algebra in terms of the normalized arithmetic wavelets as computed in Theorem 4.13 of [26] and going back in essence to [20].

Theorem A.1. *For each $n \in \mathbb{Z}$, we have the wavelet expansion*

$$\begin{aligned} L_n = e^{in\theta}\frac{\partial}{\partial\theta} &= (b_0^n + b_{+1}^n e^{i\theta} + b_{-1}^n e^{-i\theta})\frac{\partial}{\partial\theta} \\ &+ \frac{i}{4}\sum_{e\in\tau_*}\left\{n(\xi^n+\eta^n) + \frac{\eta+\xi}{\eta-\xi}\,(\xi^n-\eta^n)\right\}\bar{\vartheta}_e(\theta), \end{aligned}$$

where $e \in \tau_*$ has ideal points $\xi, \eta \in \mathbb{S}^1$ and

$$b_0^n = \begin{cases} +1, n \equiv 0(4); \\ 0, n \equiv 1(4); \\ +1, n \equiv 2(4); \\ 0, n \equiv 3(4); \end{cases} \quad b_{+1}^n = \begin{cases} 0, n \equiv 0(4); \\ +1, n \equiv 1(4); \\ -i, n \equiv 2(4); \\ 0, n \equiv 3(4); \end{cases} \quad b_{-1}^n = \begin{cases} 0, n \equiv 0(4); \\ 0, n \equiv 1(4); \\ +i, n \equiv 2(4); \\ +1, n \equiv 3(4). \end{cases}$$

This is not the simplest expression $e^{in\theta} \doteq \sum_{|e|\in\tau_*} g_n^e \bar{\vartheta}_{|e|}$, rather these particular coefficients g_n^e are specially chosen in Theorem A.1 to guarantee their suitable decay in n, cf. Theorem 6.4 of [27].

The next result, Theorem 4.11 of [26], which goes back to [24], is of basic utility and provides the Fourier expansion of the normalized arithmetic wavelets.

Theorem A.2. *If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2$, then the Fourier expansion $\bar{\vartheta}_A \sim \sum_{n\in\mathbb{Z}} c_n\ e^{in\theta}\frac{\partial}{\partial\theta}$ for $n^2 > 1$ is given by*

$$\begin{aligned} \pi i(n^3-n)\ c_n \ &= \ -[(c-a)^2+(b-d)^2]\left[\frac{(b-d)-i(a-c)}{(b-d)+i(a-c)}\right]^n \\ &+ 2(c^2+d^2)\left[\frac{d-ic}{d+ic}\right]^n \ + \ 2(a^2+b^2)\left[\frac{b-ia}{b+ia}\right]^n \\ &- [(c+a)^2+(b+d)^2]\left[\frac{(b+d)-i(a+c)}{(b+d)+i(a+c)}\right]^n, \end{aligned}$$

where the Fourier modes $c_0, c_{\pm 1}$ *are chosen to guarantee that the expansion is normalized. In particular for the mother wavelet, we have* $\bar{\vartheta} = \vartheta \sim \frac{8}{\pi i} \sum_{n \equiv 2(4)} \frac{1}{n^3 - n}\, e^{in\theta} \frac{\partial}{\partial\theta}$.

Remark A.3. A sketch of the proof is that since $\bar{\vartheta}_A$ is once continuously differentiable, we can twice integrate by parts the standard expression for the Fourier coefficients c_n with $n^2 > 1$ and derive these expressions without difficulty. For $n^2 \leqslant 1$ the corresponding equations give certain constraining equations on the zeroeth Fourier modes, which are crucial for the computation of the Kirillov-Kostant form, cf. Section 6

It is not difficult to compute directly the Fourier expansion of the (unnormalized) hyperfans as follows.

Proposition A.4. *Given* $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2$, *let*

$$\zeta_- = \frac{b - ia}{b + ia}, \quad \zeta_+ = \frac{d - ic}{d + ic} \in \mathbb{S}^1.$$

Then the hyperfan ψ_A *has Fourier expansion* $\psi_A \sim \sum c_n\, e^{in\theta}$ *with*

$$\begin{aligned} 2\pi i\; c_n = {} & \zeta_-^n \left[\frac{(d - ic)^2}{2(n+1)} \zeta_- + \frac{(d + ic)^2}{2(n-1)} \zeta_-^{-1} - \frac{c^2 + d^2}{n} \right] \\ & - \zeta_+^n \left[\frac{(d - ic)^2}{2(n+1)} \zeta_+ + \frac{(d + ic)^2}{2(n-1)} \zeta_+^{-1} - \frac{c^2 + d^2}{n} \right], \end{aligned}$$

for $n^2 > 2$, *and*

$$\begin{aligned} 2\pi i\; c_0 = {} & \frac{(d - ic)^2}{2} [\zeta_+^{-1} - \zeta_-^{-1}] - \frac{(d + ic)^2}{2} [\zeta_+ - \zeta_-] \\ & + i(c^2 + d^2)[\theta_+ - \theta_-], \\ 2\pi i\; c_{\pm 1} = {} & \mp \frac{(d \pm ic)^2}{2} [\zeta_+^{\pm 2} - \zeta_-^{\pm 2}] \pm (c^2 + d^2)[\zeta_+^{\pm 1} - \zeta_-^{\pm 1}] \\ & - \frac{i(d \mp ic)^2}{2} [\theta_+ - \theta_-], \end{aligned}$$

where $\theta_\mp = \arctan \zeta_\pm$.

Proof. ψ_A is supported on an interval with respective initial and terminal points given by ζ_-^{-1} and ζ_+^{-1}, and takes there the values

$$A^{-1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} A = \begin{pmatrix} cd & d^2 \\ -c^2 & -cd \end{pmatrix}$$

or in other words $\frac{(d-ic)^2}{2} e^{i\theta} + \frac{(d+ic)^2}{2} e^{-i\theta} - (c^2 + d^2)$. The standard expression for Fourier coefficients yields the results. □

Notice that the normalized hyperfans $\bar{\psi}_A$ have the same higher Fourier coefficients as ψ_A since $\psi_A \doteq \bar{\psi}_A$. The small Fourier modes for $\bar{\psi}_A$, and likewise those of $\bar{\theta}_A$, can be computed with some difficulty by calculating the values of these vector fields at $\frac{0}{1}, \frac{1}{1}, \frac{1}{0} \in \mathbb{S}^1$. As we do not require them, we omit the details.

Appendix B. The E_2 automorphic representation

The automorphic form that we have constructed on the universal Teichmüller space has a simple analogue on the hyperbolic plane, namely,

$$\xi(z) = E_2(z)dz,$$

where $E_2(z)$ is an almost holomorphic modular form of weight 2. Since dz transforms with respect to $\mathrm{PSL}_2(\mathbb{R})$ with the factor balancing that of $E_2(z)$, we obtain a $\mathrm{PSL}_2(\mathbb{Z})$-invariant 1-form that can be viewed as the classical counterpart of the automorphic form constructed in Section 7.

To state our results unambiguously, let us recall the following explicit formulae, c.f. [9]. Set

$$\mathbb{E}_2(z) = 1 - 24\sum_{n=1}^{\infty}\sigma(n)q^n,$$

where $\sigma(n)$ is the sum of all the positive divisors of n and we set $q = \exp(2\pi i z)$. Then $\mathbb{E}_2(z)$ is holomorphic but not quite modular invariant since

$$\mathbb{E}_2(-\frac{1}{z}) = z^2\mathbb{E}_2(z) + \frac{12}{2\pi i}z.$$

To get an honest weight 2 form, we need a correction

$$E(z) = \mathbb{E}_2(z) - \frac{3}{\pi\ \mathrm{Im}(z)},$$

but we thereby lose the holomorphicity of $\mathbb{E}_2(z)$, since

$$\frac{d}{d\bar{z}}E(z) = \frac{3i}{2\pi}\frac{1}{(\mathrm{Im}(z))^2}.$$

In the classical theory of automorphic forms, one can lift the holomorphic modular forms from the hyperbolic plane to $\mathrm{PSL}_2(\mathbb{R})$ so that they become the lowest weight vectors of the holomorphic series representations. Specifically for a weight $2k$ holomorphic form $f : \mathcal{U} \to \mathbb{C}$, where $k \geqslant 2$, we define

$$\phi_f(g) = (ci + d)^{-2k} f(g \cdot i),$$

as in [15], where $g = \left(\begin{smallmatrix} 1 & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{R})$. It follows that ϕ_f is invariant under the left action

$$\phi_f(\gamma g) = \phi_f(g), \text{ for } \gamma \in \mathrm{PSL}_2(\mathbb{Z}) \tag{B.1}$$

of $\mathrm{PSL}_2(\mathbb{Z})$. One can then define the action of $\mathrm{PSL}_2(\mathbb{R})$ on a certain space of automorphic functions generated by ϕ_f by the right action

$$(\pi(g)\phi)(h) = \phi(hg), \text{ for } h, g \in \mathrm{PSL}_2(\mathbb{R}),$$

so that ϕ_f becomes the lowest weight vector for the Lie algebra $psl_2(\mathbb{R})$ in the appropriate basis E, F, H.

To state the explicit formulae, we shall rely on explicit calculations from [11]. Let h, e, f be the standard basis of $psl_2(\mathbb{R})$ and consider the Cayley transform of this basis

$$H = \begin{pmatrix} 0 & -1 \\ i & 0 \end{pmatrix}, E = \frac{1}{2}\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, F = \frac{1}{2}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.$$

Using the standard parametrization

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{B.2}$$

of $\mathrm{PSL}_2(\mathbb{R})$, one can find an explicit action of the Lie algebra basis as follows

$$\begin{aligned} H &= -\partial_\theta, \\ E &= 2ie^{2i\theta}(y\partial_{\bar{z}} - \frac{1}{4}\partial_\theta), \\ F &= -2ie^{-2i\theta}(y\partial_{\bar{z}} - \frac{1}{4}\partial_\theta). \end{aligned}$$

Thus, one has

$$\begin{aligned} F\phi_f &= 0, \\ H\phi_f &= 2k\phi_f. \end{aligned}$$

In the case that $k = 1$, the first of these two equations no longer holds, and instead one has

$$F\phi_f = \frac{3}{\pi}.$$

It follows that ϕ_f is no longer a highest weight vector, but it also generates a one-dimensional sub-representation. Only after factorization with respect to this sub-representation does one obtain the weight 2 irreducible representation of the holomorphic series, denoted V_+.

Similarly using the complex conjugation, one gets the weight 2 representation of the antiholomorphic series denoted V_-. The one-dimensional sub-representation, denoted V_0, is common to both the weight 2 indecomposable representations, and we get a larger indecomposable representation, denoted V, which has a simple composition series

$$0 \to V_0 \to V \to V_+ \otimes V_- \to 0.$$

In order to characterize the resulting three component representation, one can use the Casimir operator, which in the explicit coordinates (B.2) has the form

$$\Delta = y^2 (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}) - y \frac{\partial^2}{\partial x \partial \theta}.$$

It annihilates the weight 2 automorphic representation generated by Φ_f and its conjugate, where $f = E_2$, to wit

$$\Delta \phi_{E_2} = 0. \tag{B.3}$$

One can check that this harmonic property (B.3) together with the automorphicity condition (B.1) and the usual growth condition of automorphic functions characterizes the three component representation V that we have obtained by lifting the E_2 modular 1-form. The problem of fundamental importance in our new theory is to construct a representation of the Lie algebra $ppsl_2(\mathbb{R})$ using a lift of our automorphic 1-form in Section 7.

The automorphic representation V that we have constructed is well known in the representation theory of $\mathrm{PSL}_2(\mathbb{R})$ as a limit of complementary series [34]. The representations V_+ and V_- are the spin 1 irreducible representations of the holomorphic and antiholomorphic discrete series. Note that the lift of Eisenstein series E_{2k}, for $k = 2, 3, \ldots$, yields spin k irreducible representations of the same series. The representation V plays a pivotal role and somehow seems to have been missed in the theory of automorphic functions for the pair $\mathrm{PSL}_2(\mathbb{Z}) \subseteq \mathrm{PSL}_2(\mathbb{R})$ Thus it is expected that the counterpart of the representation V in our new theory will play an equally fundamental role, and it is one of our main open problems to construct it explicitly.

One can ask what is an advantage of a realization of this representation as an automorphic representation. One such benefit is that we can naturally increase this representation by relaxing the automorphicity condition (B.1) from the modular group $\mathrm{PSL}_2(\mathbb{Z})$ to various subgroups including its commutant $\mathrm{PSL}_2(\mathbb{Z})'$ considered in Section 1. This immediately yields six copies of the weight 2 representation.

Finally, we note that we could also consider the weight 1 representations which correspond to the limit of the holomorphic discrete series of

$SL_2(\mathbb{R})$. The index of the commutant is $[SL_2(\mathbb{R}) : SL_2(\mathbb{R})'] = 12$ in this case, and we get correspondingly 12 copies upon relaxing the automorphicity condition.

References

[1] L. Bers, "Universal Teichmüller space", in *Analytic Methods in Mathematical Physics*, pp. 65-83, Gordon and Breach, New York, 1968.

[2] S. Bonnot, R. Penner, D. Saric, "A presentation for the baseleaf preserving mapping class group of the punctured solenoid", *Algebraic and Geometric Topology* **7**(3) (2007), 1171-1199.

[3] C. P. Boyer, J. C. Hurtubise, B. M. Mann, R. J. Milgram, "The topology of instanton moduli spaces, I: the Atiyah-Jones Conjecture", *Annals of Mathematics* **137** (1993), 561-609.

[4] F. Brown, "Multiple zeta values and periods of moduli spaces M_0^n", *Annalens Scientifiques de l'École Normale Supérieure* **42** (2009), 371-489.

[5] J. Cannon, W. Floyd, W. Parry, "Introductory remarks on Richard Thompson's groups", L'Enseignement Mathematique **42** (1996), 215-256.

[6] J. Conway, J. McKay, A. Sabbar, "On the discrete groups of Moonshine", *Proceedings of the American Mathematical Society* **132** (8) (2004), 2233-2240.

[7] J. Conway and S. Norton, "Monstrous moonshine", *Bulletin London Mathematical Society* **11** (1979), 308-339.

[8] C. Cummins, "Congruence subgroups of groups commensurable with $PSL_2(\mathbb{Z})$ of genus 0 and 1", *Experimental Mathematics* **13** (2004), 361-382.

[9] F. Diamond and J. Shurman, *A first course in modular forms*, Springer-Verlag, 2005.

[10] J. Duncan and I. Frenkel, "Rademacher sums, moonshine and gravity", *Communications in Number Theory and Physics* **5**(4) (2011), 1-128.s"

[11] P. Fleig, H. Gustaffson, A. Kleinschmidt, D. Persson, "Eisenstein series and automorphic representations", *Cambridge Studies in Advanced Mathematics* Cambridge University Press, 2018.

[12] I. Frenkel, J. Lepowsky, A. Meurman, "Vertex operator algebras and the Monster", *Pure and Applied Mathemartics* **134**, Academic Press, Boston, MA 1988.

[13] J. Fuchs, Affine Lie Algebras and Quantum Groups, Cambridge University Press (1992).

[14] L. Funar and C. Kapoudjian, "The braided Ptolemy-Thompson group is finitely presented", *Geometry and Topology* **12** (2008), 475-530.

[15] I. Gelfand, M. Graev and I. Piatetski-Shapiro, *Representation theory and automorphic functions*, Saunders Mathematics Books, Saunders, 1968.

[16] V. F. R. Jones. "Some unitary representations of Thompson?s groups F and T", *Journal of Combinatorial Algebra* **1**(1) (2014), 1-44. arXiv: 1412.7740.

[17] V. F. R. Jones. "A no-go theorem for the continuum limit of a periodic quantum spin chain", *Communications in Mathematical Physics* **357** (2018), 295-317.

[18] A. A. Kirillov, Lectures on the Orbit Method, Graduate Studies in Mathematics, Vol. 64, American Mathematical Society.

[19] G. Laget, “Groupes de Thompson projetifs de genre 0”, These de doctorat en Mathematiques, Universite Joseph-Fourier, Grenoble, France 2004.

[20] F. Malikov and R. C. Penner, “The Lie algebra of homeomorphisms of the circle”, *Advances in Mathematics* **140** (1999), 282-322.

[21] H. Minkowski, ”Zur Geometrie der Zahlen”, Verhandlungen des III. internationalen Mathematiker-Kongresses in Heidelberg, Berlin (1907), pp. 164-173, JFM 36.0281.01, archived from the original on 4 January 2015.

[22] T. J. Osborne and D. E. Steigemann, “Dynamics for holographic codes”, arXiv:1706.08823.

[23] R. C. Penner, “The decorated Teichmüller space of punctured surfaces”, *Communications in Mathematical Physics* **113** (1987), 299-339.

[24] —, “Universal constructions in Teichmüller theory”, *Advances in Mathematics* **98** (1993), 143-215.

[25] —, “Weil-Petersson volumes”, *Journal of Differential Geometry* **35** (1992), 559-608.

[26] —, *Decorated Teichmüller theory*, QGM Masters Class Series, volume 1, European Math Society (2012).

[27] —, “On Hilbert, Fourier and wavelet transforms”, *Communications on Pure and Applied Mathematics* **15** (2002), 772-814.

[28] — and D. Sáric, “Decorated Teichmüller theory of the punctured solenoid”, *Geometriae Dedicata* **132** (2008), 179-212.

[29] C. Scainci and K. Krasnov, “The Universal Phase Space of AdS$_3$ Gravity”, *Communications in Mathematical Physics* **322** (2013), 167-205.

[30] A. Pressley and G. Segal, Loop Groups, Oxford Mathematical Monographs (1986), Oxford University Press.

[31] D. Sullivan, “Linking the universalities of Milnor-Thurston, Feigenbaum and AhlforsBers”, Milnor Festschrift, Topological methods in modern mathematics (L. Goldberg and A. Phillips, eds.), Publish or Perish (1993), 543-563.

[32] A. Verjovsky and S. Nag, “Diff(S^1) and the Teichmüller spaces”, *Communications in Mathematical Physics* **130** (1990), 123-138.

[33] H. Verlinde, “Conformal field theory, two-dimensional quantum gravity and quantization of Teichmüller space”, *Nuclear Physics B* **337** (1990), 652-680.

[34] A. Vershik, I. Gelfand, M. Graev, “Representations of the group $SL(2, R)$, where R is a ring of functions”, *Uspekhi Mat. Nauk* **28** (1973), 83-128.

[35] E. Witten, “Three-dimensional quantum gravity revisited”, arXiv:0706.3359, June 2007.

Mathematics Department, Yale University, New Haven, CT 06520, USA
E-mail address: `frenkel-igor@yale.edu`

Institut des Hautes Études Scientifiques, 35 route des Chartres, Le Bois Marie, 91440 Bures-sur-Yvette, France, and Mathematics Department, UCLA, Los Angeles, CA 90095, USA
E-mail address: `rpenner@ihes.fr`

arXiv:math/9906167v2 [math.QA] 10 Aug 1999

August, 1999

Monstrous Moonshine and the Classification of CFT

(16 lectures given in Istanbul, August 1998)

Terry Gannon

Department of Mathematical Sciences, University of Alberta,
Edmonton, Alberta, Canada, T6G 2G1

e-mail: tgannon@math.ualberta.ca

Abstract

In these notes we give an introduction both to Monstrous Moonshine and to the classification of rational conformal field theories, using this as an excuse to explore several related structures and go on a little tour of modern math. We will discuss Lie algebras, modular functions, the finite simple group classification, vertex operator algebras, Fermat's Last Theorem, category theory, (generalised) Kac-Moody algebras, denominator identities, the A-D-E meta-pattern, representations of affine algebras, Galois theory, etc. This work is informal and pedagogical, and aimed mostly at grad students in math or math phys, but I hope that many interested nonexperts will find something of value here — like any good Walt Disney movie I try not to completely ignore the 'grown-ups'. My emphasis is on ideas and motivations, so these notes are intended to complement other papers and books where this material is presented with more technical detail. The level of difficulty varies significantly from topic to topic. The two parts — in fact any of the sections — can be read independently.

Table of Contents

Part 1. The classification of conformal field theories

1.1. Informal motivation 1
1.2. Lie algebras 4
1.3. Representations of finite-dimensional simple Lie algebras 10
1.4. Affine algebras and the Kac-Peterson matrices 13
1.5. The classification of physical invariants 17
1.6. The A-D-E meta-pattern 19
1.7. Simple-currents and charge-conjugation 24
1.8. Galois theory 28
1.9. The modern approach to classifying physical invariants 31

Part 2. Monstrous Moonshine

2.1. Introduction 35
2.2. Ingredient #1: Finite simple groups and the Monster 38
2.3. Ingredient #2: Modular functions and Hauptmoduls 40
2.4. The Monstrous Moonshine conjectures 46
2.5. Formal power series 48
2.6. Ingredient #3: Vertex operator algebras 51
2.7. Ingredient #4: Generalised Kac-Moody algebras 54
2.8. Ingredient #5: Denominator identities 59
2.9. Proof of the Moonshine conjectures 60

References 64

Glossary

24	§1.6
A_ℓ	$\mathrm{sl}_{\ell+1}(\mathbb{C})$
affine algebra	§1.4
algebra	§1.2
category theory	§1.1
central extension	§§1.2, 1.4
CFT	conformal field theory
character	§§1.3, 1.4
chiral algebra	§1.1
Fermat's Last Theorem	§2.3
field	§1.2
finite simple group	§2.2
genus	§§1.1, 2.3
group	§§1.3, 2.2
Hauptmodul	§2.3
highest-weight	§§1.3, 1.4
Kac-Moody algebra	§2.7
lattice	§1.6
Leech lattice Λ_{24}	§2.4
Lie algebra	§1.2
Lie group	§1.2
$\mathbb{M}$	the Monster finite simple group
manifold	§1.2
modular function	§2.3
modular group	$\mathrm{PSL}_2(\mathbb{Z})$ or $\mathrm{SL}_2(\mathbb{Z})$
P_+^k	§1.4
partition function	§§1.1, 1.2, 1.5
Perron-Frobenius	§1.7
physical invariant	RCFT partition function, §1.5
QFT	quantum field theory
RCFT	rational conformal field theory
representation	§1.3
Riemann surface	§§1.1, 2.3
S, T	Kac-Peterson matrices, (1.4.2)
subfactor theory	§1.9
topological	§§1.1, 2.3
torus	§2.3
$V^\natural$	the Moonshine module, §§2.4, 2.6
Verlinde's formula	(1.4.4)
Virasoro algebra	(1.2.7)
VOA	vertex operator algebra
Weyl group	§1.3
WZW model	§1.1

Part 1. The classification of conformal field theories

1.1. INFORMAL MOTIVATION

In this section we will sketch a very informal and 'hand-wavy' motivation to what we shall call the classification problem for rational conformal field theory (RCFT). Much of this material is more carefully treated in e.g. [18].

A CFT is a quantum field theory (QFT), usually with a two-dimensional space-time, whose symmetries include the conformal transformations. There are different approaches to CFT — for one of these see [26,27]. Another formulation which has been deeply influential is due to Graeme Segal [52]. It is motivated by string theory and is phrased in an important mathematical language called *category theory*.

A *category* consists of two types of things. One are called *objects*, and the other are called *arrows* (or *morphisms*). An arrow, written $f : A \to B$, has an initial and a final object (A and B respectively). Arrows f, g can be composed to yield a new arrow $f \circ g$, if the final object of g equals the initial object of f. Maps between categories are called *functors* if they take the objects (resp., arrows) of one to the objects (resp., arrows) of the other, and preserve composition.

The only difficulty people can have in understanding categories is in realising that there is no real content to them. It's just a language, highly abstract like the more familiar set theory, but in many contexts (a great example is the theory of knot invariants [58]) one which is both natural and suggestive. It tries to deflect some of our instinctive infatuation with objects (nouns), to the mathematically more fruitful one with structure-preserving maps between objects (verbs). A gentle introduction to the mathematics and philosophy of categories is [43]; we'll give a taste of this shortly.

The standard example of a category is called **Set**, where the 'objects' are sets, and the arrows from A to B are functions $A \to B$. A related example that Segal uses is **Vect**, where the objects are complex vector spaces and the arrows are linear maps. A rather trivial example of a functor $\mathcal{F} : \mathbf{Vect} \to \mathbf{Set}$ sends a vector space V to its underlying set, also called V — i.e. $\mathcal{F}$ simply 'forgets' the vector space structure on V and ignores the fact that the arrows f in **Vect** are linear. The other category Segal uses he calls **C**; its objects are disjoint unions of (parametrised) circles S^1, and the arrows are (conformal equivalence classes of) *cobordisms*, i.e. (Riemann) surfaces whose boundaries are those circles. Composition of arrows in **C** is accomplished by gluing the surfaces along the appropriate boundary circles.

Consider the usual definition of a one-to-one function: $f(x) = f(y)$ only when $x = y$. Category theory replaces this with the following. The arrow $f : A \to B$ is called 'monic' if for any arrow $g : C \to B$, there exists a unique arrow $h : C \to A$ such that $f \circ h = g$. So it's a sort of factorisation property. You can easily verify that in **Set** the notions of 'one-to-one' and 'monic' coincide. What does this redefinition gain us? It certainly doesn't seem any simpler. But it does change the focus from the *argument* of f, to the *global* functional behaviour of f, and a change of perspective can never be bad. And it allows us to transport the idea of one-to-one-ness to arbitrary categories. For instance, in the Riemann surface category **C**, the 'one-to-one functions' are the genus-0 cobordisms.

Or consider the notion of *product*. In category theory, we say that the triple (P, a, b) is a product of objects A, B if $a : P \to A$ and $b : P \to B$ are arrows, and if for any $f : C \to A$, $g : C \to B$, there is a unique arrow $h : C \to P$ such that $f = a \circ h$ and $g = b \circ h$. This notion unifies several constructions (each of which is the 'product' in an appropriately chosen category): Cartesian product of sets; intersection of sets; multiplication of numbers; the logical operator 'and'; direct product; infimum in a partially ordered set; etc. *Sum* can be defined similarly, unifying the constructions of disjoint union, 'or', addition, tensor product, direct sum, supremum, etc.

This generality of course comes with a price: it can wash away all of the endearing special features of a favourite theory or structure. There certainly are contexts where e.g. all human beings should be thought of as equal, but there are other contexts where the given human is none other than your mother and must be treated as such. It turns out that category theory provides a beautiful framework for understanding topological invariants such as the Jones-Reshetikhin-Turaev-Witten knot invariants (see e.g. [58]). And it seems to be a natural language for formulating CFT axiomatically, as we'll now see.

According to Segal, a CFT is a functor $\mathcal{T}$ from **C** to **Vect**, which obeys various properties. The picture comes from string theory: the fundamental object is a 'vibrating' loop; a state is given by a collection of these loops; each classical Feynman path from the initial to the final states is a world-sheet, i.e. a surface Σ whose boundary is all the loops in the initial and final states. QFT assigns a complex number I (the action) to each of these world-sheets, and the quantum amplitude, written $\langle \text{final}|\text{initial}\rangle$, will then be the integral over all worldsheets of $e^{\mathrm{i}I/\hbar}$. (The quantum amplitude is how the theory makes contact with experiment, as it tells us the probability of the given process $|\text{initial}\rangle \mapsto |\text{final}\rangle$ happening.) This is what Segal is trying to capture formally. The vector spaces in **Vect** come in in order to handle uniformly and simultaneously the various 'vibrations' of the strings. In particular there is one basic vector space H (a Hilbert space of quantum states), and the functor will take n copies of S^1 to $H^n = H \otimes \cdots \otimes H$.

The simplest interesting example here is the 'tree-level creation of a string from the vacuum'. In this case the world-sheet looks like a bowl, i.e. *topologically* is a disk D (if we imagine the bowl to be made of rubber, we could grab its rim and stretch it down flat onto the table, so we say a bowl and a disk are topologically equivalent — see also §2.3). Segal's functor gives us a linear map $\mathcal{T}(D) : \mathbb{C} \to H$ (H^0 is just $\mathbb{C}$), which we can think of equivalently as the assignment of the vector $\mathcal{T}(D)(1)$ in H to D. In the case of the standard unit disk (i.e. where the parametrisation of the boundary S^1 is simply $\theta \mapsto e^{2\pi \mathrm{i}\theta}$), this vector is cal0led the *vacuum state* $\Omega = |0\rangle$.

For another example, consider the 'vacuum-to-vacuum expectation value'. The initial and final states (objects) here are both the empty set, so the world-sheets (arrows) are closed Riemann surfaces. As usual in QFT, we can organise these by how many internal 'loops' are involved (this number is called the *genus* of the surface): topologically, 0-loop (i.e. 'tree-level') world-sheets are spheres, 1-loop world-sheets are tori, etc. These closed Riemann surfaces are discussed in more detail in §2.3. The 0-loop contribution isn't very interesting (there is only one conformal equivalence class of spheres), so let us look at the 1-loop contribution. It will be of the form $\int \mathcal{Z}([torus])\, d[torus]$, where $\mathcal{Z}$ is a complex-valued function called the partition function, and $[torus]$ is a conformal equivalence class of tori. In the Segal formalism we recover this in the following way: the functor takes $[torus]$ to a linear function from $H^0 = \mathbb{C}$ to $H^0 = \mathbb{C}$. Any such linear function is simply a 1×1 matrix, i.e. a complex number, which we call $\mathcal{Z}([torus])$.

Now there is a nice parametrisation of conformal equivalence classes of tori, as we will see more explicitly in §2.3. Namely, a representative for each class can be chosen to be of the form $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$ where τ is in the upper half plane $\mathcal{H}$. Thus we can write $\mathcal{Z}$ as a function of a complex variable τ. However, different τ correspond to the same equivalence class of tori: the redundancy is exactly captured by the modular group $\mathrm{PSL}_2(\mathbb{Z})$. Namely, τ and $\frac{a\tau+b}{c\tau+d}$ are equivalent, whenever $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. Thus $\mathcal{Z}(\tau) = \mathcal{Z}(\frac{a\tau+b}{c\tau+d})$. In other words, the partition function is modular invariant![1]

There are two sectors in CFT, a holomorphic one and an antiholomorphic one, corresponding to the two directions ('left-' and 'right-moving') of motion on a string, or the two components of the group $\mathrm{Diff}(S^1)$ of diffeomorphisms of the circle. This means that many of the quantities (e.g. the partition function) factorise into parts depending holomorphically and anti-holomorphically on the modular parameters (e.g. τ in genus 1). In a *rational* CFT there are finitely many 'primary fields' $a \in \Phi$ — the precise meaning of this is not important here, but it says that the space of states for the theory decomposes into a finite sum[2] $H = \oplus_{a,b\in\Phi} M_{ab} H_a \otimes \overline{H}_b$, where M_{ab} are nonnegative integers which count the multiplicity of $H_a \otimes \overline{H}_b$ in H. The linear maps $\mathcal{T}(\Sigma) : H^m \to H^n$ in an RCFT will factorise similarly; this 'chiral factorisation' is captured by what Segal calls the 'modular functor' [52]. The partition function becomes

$$\mathcal{Z}(\tau) = \sum_{a,b\in\Phi} M_{ab}\, \chi_a(\tau)\, \chi_b(\tau)^* \tag{1.1.1}$$

for certain holomorphic functions χ_a. One of the primary fields (we'll denote it '0') corresponds to the vacuum Ω, and uniqueness of the vacuum means that $M_{00} = 1$.

H_0 is called a *chiral algebra*; in the language of §2.6, H_0 will be a vertex operator algebra (VOA). Φ parametrises the irreducible H_0-modules and the χ's are their characters; in an RCFT we require this number to be finite. For example, for the Moonshine VOA $V^\natural$ discussed in Part 2, Φ consists of only one element.

The higher-genus behaviour of an RCFT is determined from the lower-genus behaviour, by composition of 'arrows' (i.e. the gluing together of surfaces) in **C**. See Figure 3 of [30] to find how a genus-2 surface is built up from genus-0 ones. In fact, it's generally believed that an RCFT will be uniquely determined by: (i) the choice of chiral algebra; (ii) the partition function (which tells you the spectrum of the theory, i.e. how the two sectors link up); and (iii) the structure constants C_{ab}^c, which in the Segal formalism correspond to the surfaces called 'pairs-of-pants', equivalently disks with two interior disks removed. Our approach will be to start with a chiral algebra, and find all possible partition functions. We will thus ignore the important question of existence and uniqueness of the structure constants, though at least for our chiral algebras, it seems to be generally believed that the structure constants will be unique.

---

[1] In higher-dimensional string theories, a similar argument shows more generally that *automorphic forms* will appear naturally.

[2] It seems though that 'rational' logarithmic CFT is trying to teach us the lesson that this familiar requirement can and should be weakened. See Gaberdiel-Kausch (1999).

Perhaps all chiral algebras come from standard constructions (e.g. orbifolds and the Goddard-Kent-Olive (GKO) coset construction — see e.g. [18]) involving lattices and affine Kac-Moody algebras. For instance a $\mathbb{Z}_2$-orbifold of the VOA of the Leech lattice gives us the Moonshine module $V^\natural$, and the so-called minimal models come from GKO cosets involving $A_1^{(1)}$. This is in line with the spirit of Tannaka-Krein duality (and its generalisations by Deligne and Doplicher-Roberts), which roughly says that if a bunch of things *act like* they're the set of representations of a Lie group, then they *are* the set of representations of a Lie group.

In any case, one of the simplest, best understood, and important classes (called Wess-Zumino-Witten (WZW) models — see for instance [30,59] in this volume) of RCFTs correspond to affine Kac-Moody algebras at a positive integer level k. We will have much more to say later about these algebras, but for now let us remark that Φ here will be the (finite) set P_+^k of integrable level k highest weights λ. Their chiral algebras were constructed by Frenkel and Zhu. The following sections concern the attempt to classify the partition functions corresponding to Kac-Moody algebras — see especially §1.5. I will use this theme as an excuse to describe many other things, e.g. the A-D-E meta-pattern, Lie theory, Galois, fusion rings, ... I dedicate these notes to the profound friendship developing in recent years between mathematics and physics. As Victor Kac said in his 1996 Wigner medal acceptance speech, "Some of the best ideas come to my field from the physicists. And on top of this they award me a medal. One couldn't hope for a better deal."

1.2. LIE ALGEBRAS

Lie algebras (and their nonlinear partners *Lie groups*) appear in numerous places throughout math and mathematical physics. A nice introduction is [9]; Lie theory is presented with more of a physics flavour in [24], as well as [59].

An *algebra* is a vector space with a way to multiply vectors which is compatible with the vector space structure (i.e. the vector-valued product is *bilinear*: $(a\vec{u}+a'\vec{u}')\,(b\vec{v}+b'\vec{v}') = ab\,\vec{u}\vec{v} + ab'\,\vec{u}\vec{v}' + a'b\,\vec{u}'\vec{v} + a'b'\,\vec{u}'\vec{v}'$). For example, the complex numbers $\mathbb{C}$ can be thought of as a 2-dimensional algebra over $\mathbb{R}$ (a basis is 1 and $\mathrm{i} = \sqrt{-1}$; the *scalars* here are real numbers and the *vectors* are complex numbers). The quaternions are 4-dimensional over $\mathbb{R}$ and the octonions are 8-dimensional over $\mathbb{R}$. Incidentally, these are the only finite-dimensional algebras over $\mathbb{R}$ which obey the cancellation law: $\vec{u} \neq 0$ and $\vec{u}\vec{v} = 0$ implies $\vec{v} = 0$ (the reader should try to convince himself why the familiar vector product on $\mathbb{R}^3$ fails the cancellation law). This important little fact makes several unexpected appearances in math. For instance, it is trivially possible to 'comb the hair' on the circle S^1 without 'cheating' (i.e. needing a hair-part or exploiting a bald spot): just comb the hair clockwise for example. However it is not possible to comb the hair on the sphere S^2 (e.g. your own head) without cheating. The only other k-spheres S^k which can be combed (i.e. for which there exist k linearly independent continuous vector fields) are $k = 3$ and 7. This is intimately connected with the existence of $\mathbb{C}$, the quaternions, and octonions (namely,

S^1, S^3, S^7 can be thought of as the length 1 complex numbers, quaternions, and octonions, resp.).

In a *Lie* algebra $\mathfrak{g}$, the product is usually called a 'bracket' and is written $[xy]$. It is required to be 'anti-commutative' and 'anti-associative':

$$[xy] + [yx] = 0 \tag{1.2.1a}$$

$$[x[yz]] + [y[zx]] + [z[xy]] = 0 \tag{1.2.1b}$$

(like most other equalities in math, (1.2.1b) is usually called the *Jacobi identity*). Usually we will consider Lie algebras over $\mathbb{C}$, but sometimes over $\mathbb{R}$. Note that (1.2.1a) says $[xx] = 0$.

One important consequence of bilinearity is that it is enough to know the values of all the brackets $[x_i x_j]$ for $i < j$, for any basis $\{x_1, x_2, \ldots\}$ of $\mathfrak{g}$. (The reader should convince himself of this before proceeding.)

The simplest example of a Lie algebra is $\mathfrak{g} = \mathbb{C}$ (or $\mathfrak{g} = \mathbb{R}$), with the bracket $[xy]$ identically 0. In fact, this is the only 1-dimensional Lie algebra. It is a straightforward exercise for the reader to find all 2- and 3-dimensional Lie algebras (over $\mathbb{C}$) up to isomorphism (i.e. change of basis): there are precisely 2 and 6 of them, respectively (though one of the 6 depends on a complex parameter). Over $\mathbb{R}$, there are 2 and 9 (with 2 depending on real parameters). This exercise cannot be continued much further — e.g. not all 7-dimensional Lie algebras (over $\mathbb{C}$) are known. Nor is it obvious that this would be an interesting or valuable exercise. We should suspect that our definition of Lie algebra is probably a little too general for anything obeying it to be automatically an interesting structure. More often than not, a classification turns out to be a stale and useless list.

Two of the 3-dimensional Lie algebras are important in what follows. One of them is well-known to the reader: consider the vector-product (also called cross-product) in $\mathbb{C}^3$. Taking the standard basis $\{e_1, e_2, e_3\}$ of $\mathbb{C}^3$, the bracket can be defined by the relations

$$[e_1 e_2] = e_3 \ , \qquad [e_1 e_3] = -e_2 \ , \qquad [e_2 e_3] = e_1 \ . \tag{1.2.2a}$$

This Lie algebra, denoted A_1 or $\mathrm{sl}_2(\mathbb{C})$, can be called the 'mother of all (semi-simple) Lie algebras'. A more familiar realisation of A_1 uses a basis $\{e, f, h\}$ with relations

$$[ef] = h \ , \qquad [he] = 2e \ , \qquad [hf] = -2f \ . \tag{1.2.2b}$$

The reader can find the change-of-basis (valid over $\mathbb{C}$ but not $\mathbb{R}$) showing that (1.2.2) define isomorphic *complex* (but not *real*) Lie algebras.

Another important 3-dimensional Lie algebra is called the Heisenberg algebra[3] and is the algebra of the canonical commutation relations in quantum mechanics: choosing a basis x, p, h, it is defined by

$$[xp] = h \ , \qquad [xh] = [ph] = 0 \ . \tag{1.2.3}$$

---

[3] Actually, 'Heisenberg algebra' refers to a family of Lie algebras, with (1.2.3) being the one of lowest dimension.

From our definition, it is far from clear that Lie algebras, as a class, should be natural and worth studying. After all, there are infinitely many possible axiomatic systems: why should the one defining a Lie algebra be anything special *a priori*? Perhaps this could have been anticipated by the following line of reasoning.

Axiom. Groups are important and interesting.
Axiom. Manifolds are important and interesting.

Manifolds are structures where calculus is possible; locally, a manifold looks like a piece of $\mathbb{R}^n$ (or $\mathbb{C}^n$), but these pieces can be bent and stitched together to create more interesting shapes. For instance a circle is a 1-dimensional manifold, while Einstein claimed space-time is a curved 4-dimensional one.

Definition. A Lie group is a manifold with a compatible group structure.

This means that 'multiplication' and 'inverse' are differentiable maps. $\mathbb{R}$ is a Lie group, under addition: obviously, $\mu : \mathbb{R}^2 \to \mathbb{R}$ and $\iota : \mathbb{R} \to \mathbb{R}$ defined by $\mu(a, b) = a + b$ and $\iota(a) = -a$ are both differentiable. (Why isn't $\mathbb{R}$ a Lie group under multiplication?) A circle is also a Lie group: parametrise the points with the angle θ defined mod 2π (or mod 360 if you prefer); the 'product' of the point at angle θ_1 with the point at angle θ_2 will be the point at angle $\theta_1 + \theta_2$. Surprisingly, the only other k-sphere which is a Lie group is S^3 (the product can be defined using quaternions of unit length[4], or using the matrix group $\mathrm{SU}_2(\mathbb{C})$). Many but not all Lie groups can be expressed as matrix groups. Two other examples are GL_n (invertible $n \times n$ matrices) and SL_n (ones with determinant 1).

A consequence of the above axioms is then surely:

Corollary. Lie groups should be important and interesting.

Lie group structure theory can be thought of as a major generalisation of linear algebra. The basic constructions familiar to undergraduates have important analogues valid in many Lie groups. For instance, years ago we were taught to solve linear equations and invert matrices by using elementary row operations to reduce a matrix to row-echelon form. What this says is that any matrix $A \in \mathrm{GL}_n(\mathbb{C})$ can be factorised $A = BPN$, where N is uppertriangular with 1's on the diagonal, P is a permutation matrix, and B is an uppertriangular matrix. This is essentially what is called the Bruhat decomposition of the Lie group $\mathrm{GL}_n(\mathbb{C})$. More generally (where it applies to any 'reductive' Lie group G), P will be an element of the so-called 'Weyl group' of G (of which we'll have much more to say later), and B will be in a 'Borel subgroup'.

Lie groups appear throughout physics. E.g. the orthogonal group $\mathrm{SO}_3(\mathbb{R})$ is the configuration space of a rigid body centred at the origin, while $\mathrm{SU}_2(\mathbb{C})$ is the set of states of an electron at rest. The gauge group of the Standard Model of particle physics is $\mathrm{SU}_3(\mathbb{C}) \times \mathrm{SU}_2(\mathbb{C}) \times \mathrm{U}_1(\mathbb{C})$, while the Lorentz group of special relativity is $\mathrm{SO}_{3,1}(\mathbb{R})$.

There is an important relation between Lie groups and Lie algebras.

Fact. The tangent space of a Lie group is a Lie algebra. Any (finite-dimensional real or complex) Lie algebra is the tangent space to some Lie group.

---

[4] Similarly, the 7-sphere inherits from the octonions a *nonassociative* (hence nongroup) product, compatible with its manifold structure.

More precisely, the tangent space at 1 (i.e. the set of all infinitesimal generators of the Lie group) can be given a natural Lie algebra structure. A Lie algebra, being a linearised Lie group, is much simpler and easier to handle. The Lie algebra preserves the local properties of the Lie group, though it loses global topological properties (like boundedness). A Lie group has a single Lie algebra, but a Lie algebra can correspond to many different Lie groups. The Lie algebra corresponding to both $\mathbb{R}$ and S^1 is $\mathfrak{g} = \mathbb{R}$ with trivial bracket. The Lie algebra corresponding to both $S^3 = \mathrm{SU}_2(\mathbb{C})$ and $\mathrm{SO}_3(\mathbb{R})$ is the cross-product algebra on $\mathbb{R}^3$ (usually called $\mathrm{so}_3(\mathbb{R})$). Given the above fact, a safe guess would be:

Conjecture. Lie algebras are important and interesting.

From this line of reasoning, it should be expected that historically Lie groups arose first. Indeed that is the case: the Norwegian Sophus Lie introduced them in 1873 to try to develop a Galois theory for ordinary differential equations. As the reader may be aware, Galois theory is used for instance to show that not all 5th degree (or higher) polynomials can be explicitly 'solved' using radicals — we will meet Galois theory in §1.8. Lie wanted to study the explicit solvability (integrability) of differential equations, and this led him to develop what we now call Lie theory. The importance of Lie groups however have grown well beyond this initial motivation.

An important class of Lie algebras are the so-called finite-dimensional *simple* ones. Their definition and motivation will be studied in §2.7 below, but in a certain sense they serve as building blocks for all other finite-dimensional Lie algebras.

The classification of simple finite-dimensional Lie algebras over $\mathbb{C}$ is quite important and was accomplished at the turn of the century by Killing and Cartan. There are 4 infinite families A_ℓ ($\ell \geq 1$), B_ℓ ($\ell \geq 3$), C_ℓ ($\ell \geq 2$), and D_ℓ ($\ell \geq 4$), and 5 exceptionals E_6, E_7, E_8, F_4 and G_2. A_ℓ can be thought of[5] as $\mathrm{sl}_{\ell+1}(\mathbb{C})$, the $(\ell+1) \times (\ell+1)$ matrices with trace 0. The orthogonal algebras B_ℓ and D_ℓ can be identified with $\mathrm{so}_{2\ell+1}(\mathbb{C})$ and $\mathrm{so}_{2\ell}(\mathbb{C})$, resp., where $\mathrm{so}_n(\mathbb{C})$ is all $n \times n$ anti-symmetric matrices $A^t = -A$. The symplectic algebra C_ℓ is $\mathrm{sp}_{2\ell}(\mathbb{C})$, i.e. all $2\ell \times 2\ell$ matrices A obeying $A\Omega = -\Omega A^t$, where $\Omega = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$ and I_ℓ is the identity matrix. The exceptionals can be constructed e.g. using the octonions. In all these cases the bracket is given by the commutator

$$[AB] = [A, B] := AB - BA \tag{1.2.4}$$

(it is a good exercise for the reader to confirm that the commutator satisfies (1.2.1), and that e.g. $\mathrm{sl}_n(\mathbb{C})$ is indeed closed under it). To see that (1.2.2b) truly is $\mathrm{sl}_2(\mathbb{C})$, put

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.2.5}$$

Incidentally the names A, B, C, D have no significance: since the 4 series start at $\ell = 1, 2, 3, 4$, it seemed natural to call these A, B, C, D, resp. Unfortunately a bit of bad luck happened: B_2 and C_2 are isomorphic and so at random that algebra was placed in the orthogonal series; however affine Dynkin diagrams make it clear that it really is a symplectic algebra which accidentally looks orthogonal; hence in hindsight the names of the B- and C-series really should have been switched.

---

[5] Strictly speaking these are *representations* (see next section).

This classification changes if the *field* — the choice of scalars=numbers — is changed. By a field, we mean we can add, subtract, multiply and divide, such that all the usual properties like commutativity and distributivity are obeyed. Fields will make a few different appearances in these notes. $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Q}$ are fields, while $\mathbb{Z}$ is not (you can't always divide an integer by e.g. 3, and remain in $\mathbb{Z}$). The integers mod n, which we will write $\mathbb{Z}_n$, are a field iff n is prime (the reader can verify that in e.g. $\mathbb{Z}_4$, it is not possible to divide by the field element $[2] \in \mathbb{Z}_4$ even though $[2] \neq [0]$ there). $\mathbb{C}$ and $\mathbb{R}$ are examples of fields of characteristic 0 — this means that 0 is the only integer k with the property that $kx = 0$ for all x in the field. $\mathbb{Z}_p$ is the simplest example of a field with nonzero characteristic: in $\mathbb{Z}_p$, multiplying by the integer p has the same effect as multiplying by 0, and so we say $\mathbb{Z}_p$ has characteristic p. Strange fields have important applications in e.g. coding theory and, ironically, in number theory itself — see e.g. §1.8.

As always, $\mathbb{C}$ is better behaved than e.g. $\mathbb{R}$ because every polynomial can be factorised over $\mathbb{C}$ (we say $\mathbb{C}$ is *algebraically closed*) — this implies for example that every matrix has an eigenvector over $\mathbb{C}$ but not necessarily over $\mathbb{R}$. Over $\mathbb{R}$, the difference in the simple Lie algebra classification is that each symbol $X_\ell \in \{A_\ell, \dots, G_2\}$ corresponds to a number of inequivalent algebras (over $\mathbb{C}$, each algebra has its own symbol). For example, 'A_1' corresponds to 3 different real simple Lie algebras, namely the matrix algebras $\mathrm{sl}_2(\mathbb{R})$, $\mathrm{sl}_2(\mathbb{C})$ (interpreted as a *real* vector space), and $\mathrm{so}_3(\mathbb{R})$. The simple Lie algebra classification has recently been done in any characteristic $p > 7$. It is surprising but very common that the smaller primes behave very poorly, and the classification for characteristic 2 is probably completely hopeless.

Associated with each simple algebra X_ℓ is a Weyl group, and a (Coxeter-)Dynkin diagram. The Weyl group is a finite reflection group, e.g. for A_ℓ it is the symmetric group $\mathfrak{S}_{\ell+1}$. See Figure 7 of [59] for the Weyl group of A_2. The Dynkin diagram of X_ℓ (see e.g. [24,36,38] or Figure 6 in [59]) is a graph with ℓ nodes, and with possibly some double and triple edges. It says how to construct X_ℓ abstractly using generators and relations — see §2.7. We will keep meeting both throughout these notes.

Another source of Lie algebras are the vector fields $\mathrm{Vect}(M)$ on a manifold M. A vector field v is a choice (in a smooth way) of a tangent vector $v(p) \in T_pM$ at each point of M. It can be thought of as a (1st order) differential operator, acting on functions $f : M \to \mathbb{R}$ (or $f : M \to \mathbb{C}$); at each point $p \in M$ take the directional derivative of f in the direction $v(p)$. For example the vector fields on the circle, $\mathrm{Vect}(S^1)$, can be thought of as anything of the form $g(\theta)\frac{d}{d\theta}$ where $g(\theta)$ can be any function with period 1. We can compose vector fields $u \circ v$, but this will result in a 2nd order differential operator: e.g.

$$(f(\theta)\frac{d}{d\theta}) \circ (g(\theta)\frac{d}{d\theta}) = f(\theta)\, g(\theta)\, \frac{d^2}{d\theta^2} + f(\theta)\, g'(\theta)\, \frac{d}{d\theta}\ .$$

Instead, the natural 'product' of vector fields is given by their commutator $[u, v] = u \circ v - v \circ u$, as it always results in a vector field: e.g.

$$[f(\theta)\frac{d}{d\theta}, g(\theta)\frac{d}{d\theta}] = (f(\theta)\, g'(\theta) - f'(\theta)\, g(\theta))\frac{d}{d\theta}$$

in Vect(S^1). Vect(M) with this bracket is an infinite-dimensional Lie algebra. In the case where M is a Lie group G, the Lie algebra of G can be interpreted as a certain finite-dimensional subalgebra of Vect(G) given by the 'left-invariant vector fields'.

Simple algebras need not be finite-dimensional. An example of an infinite-dimensional one is the *Witt algebra* $\mathcal{W}$, which can be defined (over $\mathbb{C}$) by the basis[6] L_n, $n \in \mathbb{Z}$, and the relations

$$[L_m L_n] = (m-n) L_{m+n} \ . \tag{1.2.6}$$

Using the realisation $L_n = -\mathrm{i}e^{-\mathrm{i}n\theta}\frac{d}{d\theta}$, the Witt algebra can also be interpreted as the polynomial subalgebra of the complexification $\mathbb{C} \otimes \mathrm{Vect}(S^1)$ — i.e. change the scalar field of Vect(S^1) from $\mathbb{R}$ to $\mathbb{C}$. Incidentally, infinite-dimensional Lie algebras need not have a corresponding Lie group: e.g. the real algebra Vect(S^1) is the Lie algebra of the Lie group $\mathrm{Diff}^+(S^1)$ of orientation-preserving diffeomorphisms $S^1 \to S^1$, but $\mathbb{C} \otimes \mathrm{Vect}(S^1)$ has no Lie group. $\mathrm{Diff}^+(S^1)$ plays a large role in CFT, by acting on the objects of Segal's category **C**.

The Witt algebra appears naturally in CFT: e.g. using the realisation $L_n = -z^{n+1}\frac{d}{dz}$ it is the polynomial subalgebra of the Lie algebra Vect($\mathbb{C}/\{0\}$). Very carelessly, Vect($\mathbb{C}/\{0\}$) is often thought of as the infinitesimal conformal transformations on a suitable neighbourhood of 0 (yet clearly L_{-2}, L_{-3}, ... are singular at 0!). Indeed the CFT literature is very sloppy when discussing the conformal group in 2-dimensions. The unfortunate fact is that, contrary to claims, *there is no infinite-dimensional conformal group* for $\mathbb{C} \cong \mathbb{R}^2$. The best we can do is the 3-dimensional group $\mathrm{PSL}_2(\mathbb{C})$ of Möbius transformations $z \mapsto \frac{az+b}{cz+d}$, which are orientation-preserving conformal transformations for the Riemann sphere $\mathbb{C} \cup \{\infty\}$. There seem to be 2 ways out of this rather embarrassing predicament. One is to argue that we are really interested in 'infinitesimal conformal invariance' in some meromorphic sense, so the full Witt algebra can appear. The other way is to argue that it is the conformal group of 'Minkowski space' $\mathbb{R}^{1,1}$ (or better, its compactification $S^1 \times S^1$) rather than $\mathbb{R}^2 \cong \mathbb{C}$ (or its compactification S^2) which is relevant for CFT. That conformal group *is* infinite-dimensional; for $S^1 \times S^1$ it consists of 2 copies of $\mathrm{Diff}^+(S^1) \times \mathrm{Diff}^+(S^1)$. For a more careful treatment of this point, see [51].

For reasons we will discuss in §1.4, we are more interested in the *Virasoro algebra* $\mathcal{V}$ rather than the Witt algebra $\mathcal{W}$. This is a '1-dimensional central extension' of $\mathcal{W}$; as a vector space $\mathcal{V} = \mathcal{W} \oplus \mathbb{C}C$ with relations given by

$$[L_m L_n] = (m-n) L_{m+n} + \delta_{n,-m} \frac{m\,(m^2-1)}{12}\, C \tag{1.2.7a}$$

$$[L_m C] = 0 \ . \tag{1.2.7b}$$

'1-dimensional central extension' means $\mathcal{V}$ has one extra basis vector C, which lies in the *centre* of $\mathcal{V}$ (i.e. $[xC] = 0$ for all $x \in \mathcal{V}$), and sending $C \to 0$ recovers $\mathcal{W}$ (i.e. takes (1.2.7a) to (1.2.6)). A common mistake in the physics literature is to regard C as a number: it is in fact a vector, though in many (but not all) representations it is mapped to a scalar multiple of the identity.

---

[6] In infinite dimensions, to avoid convergence complications, only finite linear combinations of basis vectors are generally permitted. Infinite linear combinations would involve taking some 'completion'.

The reason for the strange-looking (1.2.7a) is that we have little choice: $\mathcal{V}$ is the unique nontrivial 1-dimensional central extension of $\mathcal{W}$. The factor $\frac{1}{12}$ there is conventional but standard, and has to do with 'zeta-function regularisation' in string theory — i.e. the divergent sum $\sum_{n=1}^{\infty} n$ is 'reinterpreted' as $\zeta(-1) = \frac{-1}{12}$, where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ is the Riemann zeta function. Incidentally $\zeta(s)$ can be written as the product $\prod_p (1-p^{-s})^{-1}$ over all primes $p = 2, 3, 5, \ldots$ (try to see why); hence $\zeta(s)$ has a lot to do with primes, in particular their distribution. In fact the most famous unsolved problem in math today is the Riemann conjecture, which says that $\zeta(s) \neq 0$ whenever $\mathrm{Re}(s) \neq \frac{1}{2}$. One researcher recently described this conjecture as saying that the primes have music in them.

In CFT, L_0 is the energy operator. For example the partition function is given by $\mathcal{Z}(\tau) = \mathrm{Tr}_H(q^{L_0-c/24} q^{*\bar{L}_0-c/24})$ and the (normalised) character χ_a equals $\mathrm{Tr}_{H_a}(q^{L_0-c/24})$ for $q = e^{2\pi \mathrm{i}\tau}$. cI is the scalar multiple of the identity to which C gets sent; it has a physical interpretation [18] involving Casimir (vacuum) energy, which depends on space-time topology, and the strange shift by $c/24$ is due to an implicit mapping from the cylinder to the plane.

1.3. REPRESENTATIONS OF FINITE-DIMENSIONAL SIMPLE LIE ALGEBRAS

The representation theory of the simple Lie algebras[7] can perhaps be regarded as an enormous generalisation of trigonometry. For instance the facts that $\frac{\sin(nx)}{\sin(x)}$ can be written as a polynomial in $\cos(x)$ for any $n \in \mathbb{Z}$, and that

$$\frac{\sin(mx)\,\sin(nx)}{\sin(x)} = \sin((m+n)x) + \sin((m+n-2)x) + \cdots + \sin((m-n)x)$$

for any $m, n \in \mathbb{Z}_>$, are both easy special cases of the theory.

The classic example of an algebraic structure are the numbers, and they prejudice us into thinking that commutativity and associativity are the ideal. We have learned over the past couple of centuries that commutativity can often be dropped without losing depth and usefulness, but most interesting structures seem to obey some form of associativity. Moreover, true associativity (as opposed to e.g. anti-associativity) really simplifies the arithmetic. Given the happy 'accident' that the commutator $[x, y] := xy - yx$ in any associative algebra obeys anti-associativity, it would seem to be both tempting and natural to study the ways (if any) in which associative algebras $\mathfrak{A}$ can 'model' or *represent* a given Lie algebra. Precisely, we are looking for a map $\rho : \mathfrak{g} \to \mathfrak{A}$ which preserves the linear structure (i.e. ρ is a linear function), and which sends the bracket $[xy]$ in $\mathfrak{g}$ to the commutator $[\rho(x), \rho(y)]$ in $\mathfrak{A}$.

In practice groups (resp., algebras) often appear as symmetries (resp., infinitesimal generators of symmetries). These symmetries often act linearly. In other words, in practise the preferred associative algebras will usually be matrix algebras, and this is the usual form for a representation and the only kind we will consider. The *dimension* of these representations is the size of the matrices.

---

[7] See e.g. [25] for more details. Historically, representations of Lie algebras were considered even before representations of finite groups.

Finding all possible representations, even for the simple Lie algebras, is probably hopeless. However, it is possible to find all *finite-dimensional* representations of the simple Lie algebras, and the answer is easy to describe. Given a simple Lie algebra X_ℓ, there is a representation L_λ for each ℓ-tuple $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of nonnegative integers. λ is called a *highest-weight.* Moreover, we can take direct sums $\oplus_i L_{\lambda^{(i)}}$ of finitely many of these representations. The matrices in such a direct sum will be in block form. It turns out that, up to change-of-basis, this exhausts all finite-dimensional representations of X_ℓ.

It is common to replace 'representation ρ of $\mathfrak{g}$' with the equivalent notion of '$\mathfrak{g}$-module M' — i.e. we think of the matrices $\rho(x)$ as linear maps $M \to M$. A $\mathfrak{g}$*-module* is a vector space on which $\mathfrak{g}$ acts (on the left). Instead of considering the matrix $\rho(x)$, we consider 'products' xv (think of this as the matrix $\rho(x)$ times the column vector v) for $v \in M$. This product must be bilinear, and must obey $[xy]v = x(yv) - y(xv)$.

To get an idea of what L_λ looks like, consider A_1. Recall its generators e, f, h and relations (1.2.2b). Choose any $\lambda \in \mathbb{C}$. Define $x_0 \neq 0$ to formally obey $hx_0 = \lambda x_0$ and $ex_0 = 0$. Define inductively $x_{i+1} := fx_i$ for $i = 0, 1, \ldots$. Define M_λ to be the span of all x_i — we will see shortly that they are linearly independent (so M_λ is infinite-dimensional). M_λ is a module of A_1: the calculations $hx_{i+1} = hfx_i = ([hf] + fh)x_i = (-2f + fh)x_i$ and $ex_{i+1} = efx_i = ([ef] + fe)x_i = (h + fe)x_i$ show inductively that $hx_m = (\lambda - 2m)x_m$ and $ex_m = (\lambda - m + 1)m\, x_{m-1}$. From these the reader can show that the x_i are linearly independent. M_λ is called a *Verma module*; λ is called its highest-weight, and x_0 is called a highest-weight vector.

Now specialise to $\lambda = n \in \mathbb{Z}_\geq := \{0, 1, 2, \ldots\}$. Note that $ex_{n+1} = 0$ and $hx_{n+1} = (-n-2)x_{n+1}$. This means that, for these n, M_n contains a *submodule* with highest-weight vector x_{n+1}, isomorphic to M_{-n-2}. x_{n+1} is called a *null vector.* In other words, we could set $x_{n+1} := 0$ and still have an A_1-module. We would then get a *finite-dimensional* module which we'll call $L_n := M_n/M_{-n-2}$ (not to be confused with the Virasoro generator in (1.2.7)). Its basis is $\{x_0, x_1, \ldots, x_n\}$ and so it has dimension $n + 1$.

For example, take $n = 1$. Note that what we get in terms of the basis $\{x_0, x_1\}$ is the familiar representation $\mathrm{sl}_2(\mathbb{C})$ given in (1.2.5).

The situation for the other simple Lie algebras X_ℓ is similar.

It turns out to be hard to compare representations: ρ and ρ' could be equivalent (i.e. differ merely by a change-of-basis) but look very different. Or if we are given a representation, we may want to decompose it into the direct sum of some $L_{\lambda^{(i)}}$. When working with representations, it is often very useful to avoid much of the extraneous basis-dependent detail present in the function ρ. Finite group theory suggests how to do this: we should use *characters.* The character of an A_1-module M is given by Weyl: write M as a direct sum of eigenspaces $M(m)$ of h; then define

$$\mathrm{ch}_M(z) := \sum_m \dim\, M(m)\; e^{mz}\ , \tag{1.3.1}$$

for any $z \in \mathbb{C}$. The m are called *weights* and the $M(m)$ *weight-spaces.* For example, for

L_n the weights are $m = n, n-2, \ldots, -n$, the weight-spaces $L_n(m)$ are $\mathbb{C}\, x_{(n-m)/2}$, and

$$\mathrm{ch}_n(z) = \sum_{i=0}^{n} e^{(n-2i)z} = \frac{\sin((n+1)\, z)}{\sin(z)} \ . \tag{1.3.2}$$

Analogous formulas apply to any algebra X_ℓ: the character will then be a function of an ℓ-dimensional subspace $\mathfrak{h}$ called the *Cartan subalgebra*, spanned by all the h_i (see §2.7), so can be thought of as a complex-valued function of ℓ complex variables. The weights m will lie in the dual space to $\mathfrak{h}$ — i.e. are linear maps $\mathfrak{h} \to \mathbb{C}$ — so will have ℓ components. See for instance Figure 8 in [59]. Incidentally, ℓ is called the *rank* of X_ℓ.

Weyl's definition works: two representations are equivalent iff their characters are identical, and $M = \oplus_i L_{\lambda^{(i)}}$ iff $\mathrm{ch}_M(z) = \sum_i \mathrm{ch}_{\lambda^{(i)}}(z)$. It also is enormously simpler: e.g. the smallest nontrivial representation of E_8 is a map from $\mathbb{C}^{248}$ to the space of 248×248 matrices, while its character is a function $\mathbb{C}^8 \to \mathbb{C}$. But why is Weyl's definition natural? How did he come up with it?

To answer that question, we must remind ourselves of the characters of finite groups[8]. A representation of a finite group G is a structure-preserving map ρ (i.e. a group homomorphism) from G to matrices. The group's product becomes matrix product. In these notes we will be exclusively interested in group representations over $\mathbb{C}$. Two representations ρ, ρ' are called *equivalent* if there exists a matrix (change-of-basis) U such that $\rho'(g) = U\rho(g)U^{-1}$ for all g. The character ch_ρ is the map $G \to \mathbb{C}$ given by the trace: $\mathrm{ch}_\rho(g) = \mathrm{tr}(\rho(g))$. We see that equivalent representations will have the same character, because of the fundamental identity $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. This identity also tells us that the character is a 'class function', i.e. $\mathrm{ch}_\rho(hgh^{-1}) = \mathrm{tr}(\rho(h)\, \rho(g)\, \rho(h)^{-1}) = \mathrm{ch}_\rho(g)$ so ch_ρ is constant on each 'conjugacy class'. Group characters are also enormously simpler than representations: e.g. the smallest nontrivial representation of the Monster $\mathbb{M}$ (see Part 2) consists of almost 10^{54} matrices, each of size 196883×196883, while its character consists of 194 complex numbers. Incidentally, finite group representations behave analogously to the representations of X_ℓ: the role of the modules L_λ is played by the irreducible representations ρ_i, and any finite-dimensional representation of G can be decomposed uniquely into a direct sum of various ρ_i. The difference is that there are only finitely many ρ_i — their number equals the number of conjugacy classes of G.

We can use this group intuition here. In particular, given any Lie algebra X_ℓ and representation ρ, we can think of the map $e^x \mapsto e^{\rho(x)}$ as a representation of a Lie group $G(X_\ell)$ corresponding to X_ℓ (the exponential e^A of a matrix is defined by the usual power series; it will always converge). The trace of the matrix $e^{\rho(x)}$ will be the *group* character value at $e^x \in G(X_\ell)$, so we'll define it to be the *algebra* character value at $x \in X_\ell$. Again, it suffices to consider only representatives of each conjugacy class of $G(X_\ell)$, because the character will be a class function. Now, almost every matrix is diagonalisable (since almost any $n \times n$ matrix has n distinct eigenvalues), and so it would seem we aren't losing much by restricting $x \in X_\ell$ to *diagonalisable* matrices. Hence we may take our conjugacy class representatives to be *diagonal* matrices $x \in X_\ell$, i.e. (for $X_\ell = A_1$) to $x = zh$ for $z \in \mathbb{C}$ (h is diagonal in the x_i basis of L_λ). So the algebra character can be chosen to be a function of z. Finally, the trace of $e^{\rho(x)} = e^{z\rho(h)}$ will be given by (1.3.1). This completes the motivation for Weyl's character formula.

---

[8] Surprisingly, what we now call the characters of group representations were invented almost a decade before group representations were.

There is one other important observation we can make. Different diagonal matrices can belong to the same conjugacy class. For instance,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix} ,$$

so e^{zh} and e^{-zh} lie in the same $G(A_1) = \mathrm{SL}_2(\mathbb{C})$ conjugacy class. Hence $\mathrm{ch}_M(z) = \mathrm{ch}_M(-z)$. This symmetry $z \mapsto -z$ belongs to the *Weyl group* for A_1. Each X_ℓ has similar symmetries, and the Weyl group plays an important role in the whole theory, sort of analogous to the modular group for modular functions we'll discuss in §2.3.

Weyl found a generalisation of the right-side of (1.3.2), valid for all X_ℓ. The character of L_λ can be written as a fraction (2.8.1): the numerator will be a alternating sum over the Weyl group, and the denominator will be a product over 'positive roots'. This formula and its generalisations have profound consequences, as we'll see in §2.8.

Incidentally, the trigonometric identities given at the beginning of this section are the tensor product formula of representations (interpreted as the product and sum of characters), and the fact that an arbitrary character can be written as a polynomial in the fundamental characters, both specialised to A_1 (see (1.3.2) for the A_1 characters).

1.4. AFFINE ALGEBRAS AND THE KAC-PETERSON MATRICES

The theory of nontwisted affine Kac-Moody algebras (usually called *affine algebras* or *current algebras*) is extremely analogous to that of the finite-dimensional simple Lie algebras. Nothing infinite-dimensional tries harder to be finite-dimensional than affine algebras. Standard references for the following material are [38,41,24].

Let X_ℓ be any simple finite-dimensional Lie algebra. The affine algebra $X_\ell^{(1)}$ is essentially the *loop algebra* $\mathcal{L}(X_\ell)$, defined to be all possible 'Laurent polynomials' $\sum_{n\in\mathbb{Z}} a_n t^n$ where each $a_n \in X_\ell$ and all but finitely many $a_n = 0$. t here is an indeterminant. The bracket in $\mathcal{L}(X_\ell)$ is the obvious one: e.g. $[at^n, bt^m] = [ab]t^{n+m}$. Geometrically, $\mathcal{L}(X_\ell)$ is the Lie algebra of polynomial maps $S^1 \to X_\ell$ — hence the name (for that realisation, think of $t = e^{2\pi i\theta}$). Hence there are many generalisations of the loop algebra (e.g. any manifold in place of S^1 will do), closely related ones called *toroidal algebras* being the Lie algebra of maps $S^1 \times \cdots \times S^1 \to X_\ell$. But the loop algebra is simplest and best understood, and the only one we'll consider. Note that $\mathcal{L}(X_\ell)$ is infinite-dimensional. Its Lie groups are the *loop groups*, consisting of all loops $S^1 \to G(X_\ell)$ in a Lie group for X_ℓ.

We saw S^1 before, in the discussion of the Witt algebra. Thus the Virasoro and affine algebras should be related. In fact, the Virasoro algebra acts on the affine algebras as 'derivations', and this connection plays an important technical role in the theory.

$X_\ell^{(1)}$ is in the same relation to the loop algebra, that the Virasoro $\mathcal{V}$ is to the Witt $\mathcal{W}$. Namely, it is its (unique nontrivial 1-dimensional) central extension — see e.g. (7.7.1) of [38] for the analogue of (1.2.7a) here. In addition, for more technical reasons, a further (noncentral) 1-dimensional extension is usually made: the derivation $t\frac{d}{dt}$ is included (see footnote 33). $X_\ell^{(1)}$ is the simplest of the infinite-dimensional Kac-Moody algebras. The superscript '(1)' denotes the fact that the loop algebra was twisted by an order-1 automorphism — i.e. that it is untwisted. It is called 'affine' because of its Weyl group, as we shall see.

Central extensions are a common theme in today's infinite-dimensional Lie theory[9]. Their *raison d'être* is always the same: a richer supply of representations. For example, $\mathcal{W}$ has several representations, but no nontrivial one is an 'irreducible unitary positive-energy representation' — the kind of greatest interest in math phys. On the other hand, its central extension $\mathcal{V}$ has a rich supply of those representations (e.g. there's one for each choice of $c > 1, h > 0$, namely the Verma module $V_{c,h}$ corresponding to $L_0x_0 = hx_0, Cx_0 = cx_0$). At the level of groups, central extensions allow *projective* representations (i.e. representations up to a scalar factor) to become true representations. Projective representations (hence central extensions) appear naturally in QFT because a quantum state vector $|v\rangle$ is physically indistinguishable from any nonzero scalar multiple $\alpha|v\rangle$.

All of the quantities associated to X_ℓ have an analogue here: Dynkin diagram, Weyl group, weights,... For instance, the affine Dynkin diagram is obtained from the Dynkin diagram for X_ℓ by adding one node. See for example Figure 9 of [59]. The extra node is always labelled by a '0'. The Cartan subalgebra $\mathfrak{h}$ here will be $(\ell+2)$-dimensional. Many of these details will be discussed in more detail in §2.7 below.

The construction of $X_\ell^{(1)}$ is so trivial that it seems surprising anything interesting and new can happen here. But a certain 'miracle' happens...

No interesting representation of $X_\ell^{(1)}$ is finite-dimensional. The analogue for $X_\ell^{(1)}$ of the finite-dimensional representations of X_ℓ are called the *integrable highest-weight representations*, and will be denoted L_λ. The *highest-weight* λ here will be an $(\ell+1)$-tuple $(\lambda_0, \lambda_1, \dots, \lambda_\ell)$, $\lambda_i \in \mathbb{Z}_\geq$ (strictly speaking, it will be an $(\ell+2)$-tuple, but the extra component is not important and is usually ignored). As for X_ℓ, the highest-weights can be thought of as the assignment of a nonnegative integer to each node of the Dynkin diagram. The construction of L_λ is as in the finite-dimensional case. They are called *integrable* because they are precisely those highest-weight representations which can be 'integrated' to a projective representation of the corresponding loop group, and hence a representation of a central extension of the loop group.

We define the character χ_λ as in (1.3.1), though now the weights m will be $(\ell+2)$-tuples, and there will be infinitely many of them. χ_λ will be a complex-valued function of $\ell+2$ complex variables $(\vec{z}, \tau, u)$ (see (1.4.1a) below). It can be written as an alternating sum over the Weyl group W, over a 'nice' denominator. The difference here is that W is now infinite.

Perhaps most of the interest in affine algebras can be traced to the 'miracle' that their Weyl groups are a semidirect product $Q^\vee \rtimes \overline{W}$ of translations in a lattice $\mathbb{Q}^\vee$ (the ℓ-dimensional 'co-root lattice' of X_ℓ — see §1.6) with the (finite) Weyl group $\overline{W}$ of X_ℓ.

---

[9] Incidentally the *finite-dimensional* simple Lie algebras do not have nontrivial central extensions.

See Figure 10 of [59] for the Weyl group of $A_2^{(1)}$. 'Semidirect product'[10] means that any element of W can be written uniquely as (t, w) for some translation t and some $w \in \overline{W}$, and $(t, w) \circ (t', w') = (\text{stuff}, w \circ w')$.

One thing this implies is that χ_λ will be of the form 'theta function'/denominator. Theta functions are classically-studied modular forms (we will discuss these terms in §2.3), and thus the modular group $\mathrm{SL}_2(\mathbb{Z})$ will make an appearance! To make this more precise, consider the highest-weight $\lambda = (\lambda_0, \lambda_1)$ of $A_1^{(1)}$, and write $k = \lambda_0 + \lambda_1$. Then

$$\chi_\lambda = \frac{\Theta^{(k+2)}_{\lambda_1+1} - \Theta^{(k+2)}_{-\lambda_1-1}}{\Theta^{(2)}_1 - \Theta^{(2)}_{-1}} \tag{1.4.1a}$$

where these functions all depend on 3 complex variables z, τ, u, and

$$\Theta^{(n)}_m(z, \tau, u) := e^{-2\pi i n u} \sum_{\ell \in \mathbb{Z} + \frac{m}{2n}} \exp[\pi i n \tau \ell^2 - 2\sqrt{2}\pi i n \ell z] \ . \tag{1.4.1b}$$

In (1.4.1a) we can see the alternating sum over the Weyl group of A_1 in the numerator (and denominator, since we've used the $A_1^{(1)}$ denominator identity in writing (1.4.1a)). For general $X_\ell^{(1)}$, the denominator will always be independent of λ, and the theta function (1.4.1b) will become a multidimensional one involving a sum over $Q^\vee$ shifted by some weight and appropriately rescaled. The (co-)root lattice of A_1 is $\sqrt{2}\mathbb{Z}$. The key variable in (1.4.1a) is the modular one τ, which will lie in the upper half complex plane $\mathcal{H}$ (in order to have convergence). In the applications to CFT, the other variables are often set to 0.

The number k introduced in (1.4.1a) plays an important role in the general theory. In the representation L_λ, the central term C will get sent to some multiple of the identity — the multiplier is labelled k and is called the *level* of the representation. For any $X_\ell^{(1)}$ there is a simple formula expressing the level k in terms of the highest-weight λ; e.g. for $A_\ell^{(1)}$ and $C_\ell^{(1)}$ it is given by $k = \lambda_0 + \lambda_1 + \cdots + \lambda_\ell$. Write P_+^k for the (finite) set of level k highest-weights (so the size of P_+^k for $A_\ell^{(1)}$ is $\binom{k+\ell}{\ell}$). An important weight in P_+^k is $(k, 0, \ldots, 0)$. We will denote this '0'. In RCFT it corresponds to the vacuum.

The modular group $\mathrm{SL}_2(\mathbb{Z})$ acts on the Cartan subalgebra $\mathfrak{h}$ of $X_\ell^{(1)}$ in the following way:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (\vec{z}, \tau, u) = (\frac{\vec{z}}{c\tau + d}, \frac{a\tau + b}{c\tau + d}, u - \frac{c\,\vec{z}\cdot\vec{z}}{2(c\tau + d)})$$

Under this action, the characters χ_λ also transform nicely: in particular we find for any level k weight λ

$$\chi_\lambda(\frac{\vec{z}}{\tau}, \frac{-1}{\tau}, u - \frac{\vec{z}\cdot\vec{z}}{2\tau}) = \sum_{\mu \in P_+^k} S_{\lambda\mu}\, \chi_\mu(\vec{z}, \tau, u) \tag{1.4.2a}$$

$$\chi_\lambda(\vec{z}, \tau + 1, u) = \sum_{\mu \in P_+^k} T_{\lambda\mu}\, \chi_\mu(\vec{z}, \tau, u) \tag{1.4.2b}$$

---

[10] This is also discussed briefly in section 2.2.

where S and T are complex matrices called the *Kac-Peterson matrices*. S will always be symmetric and unitary, and has many remarkable properties as we shall see. Its entries are related to Lie group characters at elements of finite order (see (1.4.5) below). T is diagonal and unitary; its entries are related to the eigenvalues of the quadratic Casimir.

For example, consider $A_1^{(1)}$ at level k. Then S and T will be $(k+1)\times(k+1)$ matrices given by

$$S_{\lambda\mu}=\sqrt{\frac{2}{k+2}}\,\sin(\pi\,\frac{(\lambda_1+1)(\mu_1+1)}{k+2})\ ,\qquad T_{\lambda\mu}=\exp[\pi\mathrm{i}\,\frac{(\lambda_1+1)^2}{2(k+2)}-\frac{\pi\mathrm{i}}{4}]\,\delta_{\lambda,\mu}\ . \tag{1.4.3}$$

One important place S appears is the famous *Verlinde formula*

$$N_{\lambda\mu}^{\nu}=\sum_{\kappa\in P_+^k}\frac{S_{\lambda\kappa}\,S_{\mu\kappa}\,S_{\nu\kappa}^*}{S_{0,\kappa}} \tag{1.4.4}$$

for the fusion coefficients $N_{\lambda\mu}^{\nu}$ of the corresponding RCFT. We will investigate some consequences of this formula in a later section. The fusion coefficients for the affine algebras are well-understood; see e.g. Section 4 of [59] for their interpretation (usually called the Kac-Walton formula) as 'folded tensor product coefficients'.

We will see in §1.7 that symmetries of the extended Dynkin diagram have consequences for S and T (simple-currents, charge-conjugation). There is a 'Galois action' on S which we will discuss in §1.8. There is a strange property of S and T called *rank-level duality* (see e.g. [45]): the matrices for $A_\ell^{(1)}$ at level k are closely related to those of $A_{k-1}^{(1)}$ at level $\ell+1$, and similar statements hold for $B_\ell^{(1)}$, $C_\ell^{(1)}$ and $D_\ell^{(1)}$. Another reason S is mathematically interesting is the formula

$$\frac{S_{\lambda\mu}}{S_{0\mu}}=\mathrm{ch}_{\bar\lambda}(-2\pi\mathrm{i}\,\frac{\overline{\mu+\rho}}{k+h^\vee})\ . \tag{1.4.5}$$

The right-side is a character of X_ℓ, and $\bar\lambda=(\lambda_1,\dots,\lambda_\ell)$ means ignore the extended node. ρ is the 'Weyl vector' $(1,1,\dots,1)$ and $h^\vee$ is called the *dual Coxeter number* and is the level of ρ. For $A_\ell^{(1)}$, $h^\vee=\ell+1$. Of course the right-side can also be regarded as a character for a Lie group associated to X_ℓ, in which case the argument would have to be exponentiated and would correspond to an element of finite order in the group. These numbers (1.4.5) have been studied by many people (most extensively by Pianzola) and have some nice properties. For instance Moody-Patera (1984) have argued that exploiting them leads to some quick algorithms for computing e.g. tensor product coefficients. Kac [37] found a curious application for them: a Lie theoretic proof of 'quadratic reciprocity'!

Quadratic reciprocity is one of the gems of classical number theory. It tells us that the equations

$$\begin{aligned}x^2&\equiv a\ (\mathrm{mod}\ b)\\ y^2&\equiv b\ (\mathrm{mod}\ a)\end{aligned}$$

are related; more precisely, for fixed a and b (for simplicity take them to both be primes $\neq 2$) the questions of whether there is a solution x to the first equation and a solution y to the second, are related. They will both have the same yes or no answer, unless $a \equiv b \equiv 3 \pmod 4$, in which case they will have opposite answers. E.g. take $a = 23$ and $b = 3$, then we know the first equation does not have a solution (since $a \equiv 2 \pmod 3$ and $x^2 \equiv 2$ doesn't have a solution mod 3), and hence the second equation must have a solution (indeed, $y = 7$ works). There are now many proofs for quadratic reciprocity, and Kac used Lie characters at elements of finite order to find another one.

What is interesting here is that Kac's proof uses only certain special weights for A_ℓ. The natural question is: is it possible to find any generalisations of quadratic reciprocity using other weights and algebras? Many generalisations of quadratic reciprocity are known; will generalising Kac's argument recover them, or will they perhaps yield new reciprocity laws? It seems no one knows.

The relation (1.4.5) is important because it connects finite-dimensional Lie data with infinite-dimensional Lie data. The 'conceptual arrow' can be exploited both ways: in the generalisations of the arguments of §1.9 to other algebras, (1.4.5) allows us to use our extensive knowledge of finite-dimensional algebras to squeeze out some information in the affine setting; but also it is possible to use the richer symmetries of the affine data to see 'hidden' symmetries in finite-dimensional data. For example it can be used (Gannon-Walton 1995) to find a sort of Galois symmetry of dominant weight multiplicities in X_ℓ, which would be difficult or impossible to anticipate without (1.4.5).

1.5. The classification of physical invariants

We are interested in the following classification problem. Choose any affine algebra $X_\ell^{(1)}$ and level $k \in \mathbb{Z}_\geq$. Find all matrices $M = (M_{\lambda\mu})_{\lambda,\mu \in P_+^k}$ such that

(P1) $MS = SM$ and $MT = TM$, where S, T are the Kac-Peterson matrices (1.4.2);
(P2) each entry $M_{\lambda\mu} \in \mathbb{Z}_\geq$;
(P3) $M_{00} = 1$.

Any such M, or equivalently the corresponding partition function $\mathcal{Z} = \sum_{\lambda,\mu} M_{\lambda\mu}\chi_\lambda\chi_\mu^*$, is called a *physical invariant*.

The first and most important classification of physical invariants was the Cappelli-Itzykson-Zuber A-D-E classification for $A_1^{(1)}$ at all levels k [8]. We will give their result shortly. This implies for instance the minimal model RCFT classification, as well as the $N = 1$ super(symmetric)conformal minimal models. The other classifications of comparable magnitude are $A_2^{(1)}$ for all k; $A_\ell^{(1)}$, $B_\ell^{(1)}$ and $D_\ell^{(1)}$ for all $k \leq 3$; $(A_1 \oplus A_1)^{(1)}$ for all levels (k_1, k_2); and $(u(1) \oplus \cdots \oplus u(1))^{(1)}$ for all (matrix-valued) levels k. See e.g. [29] for references. The most difficult of these classifications is for $A_2^{(1)}$, done by Gannon (1994).

In other words, very little in this direction has been accomplished in the 15 or so years this problem has existed. But this is not really a good measure of progress. The effort instead has been directed primarily towards the full classification; most of these partial results are merely easy spin-offs from that more serious and ambitious assault.

The proof in [8] was very complicated and followed the following lines. First, an explicit basis was found for the vector space (called the 'commutant') of all matrices obeying (P1). Then (P2) and (P3) were imposed. Unfortunately their proof was long and formidable. Others tried to apply their approach to $A_2^{(1)}$, but without success. The eventual proof for $A_2^{(1)}$ was completely independent of the [8] argument, and exploited more of the structure implicit in the problem. As the $A_2^{(1)}$ argument became more refined, it became the model for the general assault. In §1.9 we sketch this new approach.

From this more general perspective, of these completed classifications only the level 2 $B_\ell^{(1)}$ and $D_\ell^{(1)}$ ones will have any lasting value (the orthogonal algebras at level 2 behave very peculiarly, possess large numbers of exceptional physical invariants, and must be treated separately). The others behave more generically and will fall out as special cases once the more general classifications are concluded. Other classifications which should be straightforward with our present understanding are $C_2^{(1)}$ at all k; $G_2^{(1)}$ at all k; and $B_\ell^{(1)}$ and $D_\ell^{(1)}$ at $k = 4$. The $C_2^{(1)}$ should be easiest and would imply the $C_\ell^{(1)}$ level 2, as well as the $B_\ell^{(1)}$ and $D_\ell^{(1)}$ level 5, classifications. A very safe conjecture is that the only *exceptional* physical invariants (we define this term in §1.7) for $C_2^{(1)}$ occur at $k = 3, 7, 8, 12$ — this is known to be true for all $k \leq 500$. $G_2^{(1)}$ would be more difficult but also much more valuable; its only known exceptionals occur at $k = 3, 4$, and these are the only exceptionals for $k \leq 500$, and a very safe conjecture is that there are no other $G_2^{(1)}$ exceptionals. $B_\ell^{(1)}$ and $D_\ell^{(1)}$ at level 4 will also be more difficult, but also would be valuable; less is understood about its physical invariants and there is a good chance new exceptionals exist there.

The most surprising thing about the known physical invariant classifications is that there so few surprises: almost every physical invariant is 'generic'. We will see that the symmetries of the extended Dynkin diagram give rise to general families of physical invariants. We will call any physical invariants which do not arise in these generic ways (i.e. using what are called simple-currents or conjugations), *exceptional*. Many exceptionals have been found, and now we are almost at the point where we can safely conjecture the complete list of physical invariants for $X_\ell^{(1)}$ at any k, for X_ℓ a simple algebra.

Unfortunately the classification for semi-simple algebras $X_{\ell_1} \oplus \cdots \oplus X_{\ell_s}$ does not reduce to the one for simple ones. In fact, *any explicit classification of the physical invariants for $X^{(1)}$, for all semi-simple X, would easily be one of the greatest accomplishments in the history of math,* for it would include as a small part such monumental things as an explicit classification of all positive-definite integral lattices. Thus we unfortunately cannot expect an *explicit* classification for the *semi*-simple algebras.

To make this discussion more concrete and explicit, consider $A_1^{(1)}$. For convenience drop λ_0, so $P_+^k = \{0, 1, \ldots, k\}$. Write J for the permutation (called a simple-current) $Ja := k - a$. Then the complete list of physical invariants for $A_1^{(1)}$ is

$$\mathcal{A}_{k+1} = \sum_{a=0}^{k} |\chi_a|^2 \ , \qquad \text{for all } k \geq 1$$

$$\mathcal{D}_{\frac{k}{2}+2} = \sum_{a=0}^{k} \chi_a \, \chi_{J^a a}^* \ , \qquad \text{whenever } \frac{k}{2} \text{ is odd}$$

$$\begin{aligned}
\mathcal{D}_{\frac{k}{2}+2} &= |\chi_0+\chi_{J0}|^2+|\chi_2+\chi_{J2}|^2+\cdots+2|\chi_{\frac{k}{2}}|^2\ , && \text{whenever } \frac{k}{2} \text{ is even}\\
\mathcal{E}_6 &= |\chi_0+\chi_6|^2+|\chi_3+\chi_7|^2+|\chi_4+\chi_{10}|^2\ , && \text{for } k=10\\
\mathcal{E}_7 &= |\chi_0+\chi_{16}|^2+|\chi_4+\chi_{12}|^2+|\chi_6+\chi_{10}|^2\\
&\quad +\chi_8\,(\chi_2+\chi_{14})^*+(\chi_2+\chi_{14})\,\chi_8^*+|\chi_8|^2\ , && \text{for } k=16\\
\mathcal{E}_8 &= |\chi_0+\chi_{10}+\chi_{18}+\chi_{28}|^2+|\chi_6+\chi_{12}+\chi_{16}+\chi_{22}|^2\ , && \text{for } k=28\ .
\end{aligned}$$

The physical invariants $\mathcal{A}_n$ and $\mathcal{D}_n$ are generic, corresponding respectively to the order 1 (i.e. identity) and order 2 (i.e. the simple-current J) Dynkin diagram symmetries, as we shall see in §1.7. Physically, they are the partition functions of WZW models on $\mathrm{SU}_2(\mathbb{C})$ and $\mathrm{SO}_3(\mathbb{R})$ group manifolds, resp. The exceptionals $\mathcal{E}_6$ and $\mathcal{E}_8$ are best interpreted as due to the $C_{2,1}\supset A_{1,10}$ and $G_{2,1}\supset A_{1,28}$ conformal embeddings (see §1.7; standard notation is to write '$X_{\ell,k}$' for '$X_\ell^{(1)}$ and level k'). The $\mathcal{E}_7$ exceptional is harder to interpret, but can be thought of as the first in an infinite series of exceptionals involving rank-level duality and D_4 triality.

Around Christmas 1985, Zuber wrote Kac about the $A_1^{(1)}$ physical invariant problem, and mentioned the physical invariants he and Itzykson knew at that point (what we now call $\mathcal{A}_\star$ and $\mathcal{D}_{even}$). A few weeks later, Kac wrote back saying he found one more invariant, and jokingly pointed out that it must be indeed quite exceptional as the exponents of E_6 appeared in it. "I must confess that I didn't pay much attention to that last remark (I hardly knew what Coxeter exponents were, at the time!)" [63]. By spring 1986, Cappelli arrived in Paris and got things moving again; together Cappelli-Itzykson-Zuber found $\mathcal{E}_7$, $\mathcal{D}_{odd}$, and then $\mathcal{E}_8$, and struggled to find more. "And it is only in August [1986], during a conversation with Pasquier, in which he was showing me his construction of lattice models based on Dynkin diagrams, that I suddenly remembered this cryptic but crucial! observation of Victor, rushed to the library to find a list of the exponents of the other algebras... and found with the delight that you can imagine that they were matching our list" [63]. Thus the A-D-E pattern to these physical invariants was discovered.

1.6. THE A-D-E META-PATTERN

Before we discuss *meta-patterns* in math, let's introduce the notion of *lattice*[11], a simple geometric structure we'll keep returning to in these notes. The standard reference for lattice theory is [13].

Consider the real vector space $\mathbb{R}^{m,n}$: its vectors look like $\vec{x}=(\vec{x}_+;\vec{x}_-)$ where $\vec{x}_+$ and $\vec{x}_-$ are m- and n-component vectors respectively, and dot products are given by $\vec{x}\cdot\vec{y}=\vec{x}_+\cdot\vec{y}_+-\vec{x}_-\cdot\vec{y}_-$. The dot products $\vec{x}_\pm\cdot\vec{y}_\pm$ are given by the usual product and sum of components. For example, the familiar Euclidean (positive-definite) space is $\mathbb{R}^n=\mathbb{R}^{n,0}$, while Minkowski space is $\mathbb{R}^{3,1}$.

---

[11] There are many words in math which have several incompatible meanings. For example, there are *vector* fields and *number* fields, and modular *forms* and modular *representations*. 'Lattice' is another of these words. Aside from the geometric meaning we will use, it also refers to a 'partially ordered set'.

Now choose any basis $B = \{\vec{x}_1, \ldots, \vec{x}_{m+n}\}$ in $\mathbb{R}^{m,n}$. So $\mathbb{R}^{m,n} = \mathbb{R}\vec{x}_1 + \cdots + \mathbb{R}\vec{x}_{m+n}$. Define the set $\Lambda(B) := \mathbb{Z}\vec{x}_1 + \cdots + \mathbb{Z}\vec{x}_{m+n}$. This is a *lattice*, and all lattices can be formed in this way[12]. So a lattice is discrete and is closed under sums and integer multiples. For example, $\mathbb{Z}^{m,n}$ is a lattice (take the standard basis in $\mathbb{R}^{m,n}$). A more interesting lattice is the hexagonal lattice (also called A_2), given by the basis $B = \{(\frac{\sqrt{2}}{2}, \frac{\sqrt{6}}{2}), (\sqrt{2}, 0)\}$ of $\mathbb{R}^2$ — try to plot several points. If you wanted to slide a bunch of coins on a table together as tightly as possible, their centres would form this hexagonal lattice. Another important lattice is $II_{1,1} \subset \mathbb{R}^{1,1}$, given by $B = \{(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}; \frac{-1}{\sqrt{2}})\}$; equivalently it can be thought of as the set of all pairs $(a, b) \in \mathbb{Z}^2$ with dot product

$$(a, b) \cdot (c, d) = ad + bc\ . \tag{1.6.1}$$

It is important to note that different choices of basis may or may not result in a different lattice. For a trivial example, consider $B = \{1\}$ and $B' = \{-1\}$ in $\mathbb{R} = \mathbb{R}^{1,0}$: they both give the lattice $\mathbb{Z} = \mathbb{Z}^{1,0}$. Two lattices are called *equivalent* if they only differ by a change-of-basis. E.g. $B = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})\}$ in $\mathbb{R}^2$ yields a lattice equivalent to $\mathbb{Z}^2$.

The *dimension* of the lattice is $m + n$. The lattice is called *positive-definite* if it lies in some $\mathbb{R}^m$ (i.e. $n = 0$). The lattice is called *integral* if all dot products $\vec{x} \cdot \vec{y}$ are integers, for $\vec{x}, \vec{y} \in \Lambda$. A lattice Λ is called *even* if it is integral and in addition all norms $\vec{x} \cdot \vec{x}$ are *even* integers. For example, $\mathbb{Z}^{m,n}$ is integral but not even, while A_2 and $II_{1,1}$ are even. The *dual* Λ^* of a lattice Λ consists of all vectors $\vec{x} \in \mathbb{R}^{m,n}$ such that $\vec{x} \cdot \Lambda \subset \mathbb{Z}$. So a lattice is integral iff $\Lambda \subseteq \Lambda^*$. A lattice is called *self-dual* if $\Lambda = \Lambda^*$. $\mathbb{Z}^{m,n}$ and $II_{1,1}$ are self-dual but A_2 is not.

There are lots of ‘meta-patterns’ in math, i.e. collections of seemingly different problems which have similar answers. Once one of these meta-patterns is identified it is always helpful to understand what is responsible for it. For example, while I was writing up my PhD thesis I noticed in several places the numbers 1, 2, 3, 4, and 6. For instance $\cos(2\pi r) \in \mathbb{Q}$ for $r \in \mathbb{Q}$ iff the denominator of r is 1, 2, 3, 4, or 6. This pattern was easy to explain: they are precisely those positive integers n with Euler totient $\phi(n) \leq 2$, i.e. there are at most 2 positive numbers less than n coprime[13] to n. The other incidences of these numbers can usually be reduced to this $\phi(n) \leq 2$ property (e.g. the dimension of the number field $\mathbb{Q}[\cos(2\pi\frac{a}{b})]$ (see §1.8) considered as a vector space over $\mathbb{Q}$ will be $\phi(b)/2$).

A more interesting meta-pattern involves the number 24 and its divisors. One sees 24 wherever modular forms naturally appear. For instance, we see it in the critical dimensions in string theory: 24 + 2 and 8 + 2. Another example: the dimensions of even self-dual positive-definite lattices must be a multiple of 8 (e.g. the E_8 root lattice defined shortly has dimension 8, while the Leech lattice discussed in §2.4 has dimension 24). The meta-pattern 24 is also understood: the fundamental problem for which it is the answer is the following one. Fix n, and consider the congruence $x^2 \equiv 1 \pmod{n}$. Certainly in order to have a chance of satisfying this, x and n must be coprime. The extreme situation is when *every* number x coprime to n satisfies this congruence:

---

[12] In most presentations a lattice is permitted to have smaller dimension than its ambient space, however that freedom gains no real generality.

[13] We say m,n are *coprime* if any prime p which divides m does not divide n, and vice versa.

$$\gcd(x,n)=1 \qquad \Longleftrightarrow \qquad x^2 \equiv 1 \pmod{n}\ . \tag{1.6.2}$$

The reader can try to verify the following simple fact: n obeys this extreme situation (1.6.2) iff n divides 24.

What does this congruence property have to do with these other occurrences of 24? Let Λ be an even self-dual positive-definite lattice of dimension n. Then an elementary argument shows that there will exist an n-tuple $\vec{a}=(a_1,\ldots,a_n)$ of odd integers with the property that 8 must divide $\vec{a}\cdot\vec{a}=\sum_i a_i^2$. But $a_i^2 \equiv 1 \pmod 8$, and so we get $8|n$.

A much deeper and still not-completely-understood meta-pattern is called A-D-E (see [1] for a discussion and examples). The name comes from the so-called *simply-laced algebras*, i.e. the simple finite-dimensional Lie algebras whose Dynkin diagrams — see Figure 6 in [59] — contain only single edges (i.e. no arrows). These are the $A_\star$- and $D_\star$-series, along with the E_6, E_7 and E_8 exceptionals. The claim is that many other problems, which don't seem to have anything directly in common with simple Lie algebras, have a solution which falls into this A-D-E pattern (for an object to be meaningfully labelled X_ℓ, some of the data associated to the algebra X_ℓ should reappear in some form in that object). Let's look at some examples.

Consider even positive-definite lattices Λ. The smallest possible nonzero norm in Λ will be 2, and the vectors of norm 2 are special and are called *roots*. The reason they are special is that reflecting through them will always be an automorphism of Λ. That is, the reflection $\vec{u}\mapsto\vec{u}-2\frac{\vec{u}\cdot\vec{\alpha}}{\vec{\alpha}\cdot\vec{\alpha}}\,\vec{\alpha}$ through $\vec{\alpha}\neq\vec{0}$ won't in general map Λ to itself, unless $\vec{\alpha}$ is a root of Λ. It is important in lattice theory to know the lattices which are spanned by their roots; it turns out these are precisely the orthogonal direct sums of lattices called A_n, D_n, and E_6, E_7 and E_8. They carry those names for a number of reasons. For example, the lattice called X_n will have a basis $\{\vec{\alpha}_1,\ldots,\vec{\alpha}_n\}$ with the property that the matrix $A_{ij}:=\vec{\alpha}_i\cdot\vec{\alpha}_j$ is the Cartan matrix (see §2.7) for the Lie algebra X_n! Also, the reflection group generated by reflections in the roots of the lattice X_n will be isomorphic to the Weyl group of the Lie algebra X_n. Finally, to any simple Lie algebra there is canonically associated a lattice called the root lattice; for the simply-laced algebras, these will equal the corresponding lattice of the same name. Incidentally, the root lattices for the non-simply-laced simple algebras will (up to rescalings) be direct sums of the simply-laced root lattices.

We have already met the A_2 lattice: it is the densest packing of circles in the plane. It has long been believed that the obvious pyramidal way to pack oranges is also the densest possible way — the centres of the oranges form the A_3 root lattice. A controversial proof for this famous conjecture has been offered by W.-Y. Hsiang in 1991; in 1998 a new proof by Hale *et al* has been proposed. The densest known packings in dimensions 4,5,6,7,8 are D_4, D_5, E_6, E_7, E_8, resp. E_8 is the smallest even self-dual positive-definite lattice.

A famous A-D-E example is called the McKay[14] correspondence. Consider any finite subgroup G of the Lie group $\mathrm{SU}_2(\mathbb{C})$ (i.e. the 2×2 unitary matrices with determinant 1). For example, there is the cyclic group $\mathbb{Z}_n$ of n elements generated by the matrix

$$M_n=\begin{pmatrix}\exp[2\pi\mathrm{i}/n] & 0\\ 0 & \exp[-2\pi\mathrm{i}/n]\end{pmatrix}$$

---

[14] He is the same John McKay we will celebrate in section 2.1.

Let R_i be the irreducible representations of G. For instance, for $\mathbb{Z}_n$, there are precisely n of these, all 1-dimensional, given by sending the generator M_n to $\exp[2k\pi\mathrm{i}/n]$ for each $k = 1, 2, \ldots, n$. Now consider the tensor product $G \otimes R_i$, where we interpret $G \subset \mathrm{SU}_2(\mathbb{C})$ here as a 2-dimensional representation. We can decompose that product into a direct sum $\oplus_j m_{ij} R_j$ of irreducibles (the m_{ij} here are multiplicities). Now create a graph with one node for each R_i, and with the ith and jth nodes ($i \neq j$) connected with precisely m_{ij} directed edges $i \to j$. If $m_{ij} = m_{ji}$, we agree to erase the double arrows from the m_{ij} edges. Then McKay observed that the graph of any G will be a distinct extended Dynkin diagram of A-D-E type! For instance, the cyclic group with n elements corresponds to the extended graph of A_{n-1}.

How was McKay led to his remarkable correspondence? He knew that the sum of the 'marks' $a_i = 1, 2, 3, 4, 5, 6, 4, 2, 3$ associated to each node of the extended E_8 Dynkin diagram equaled 30, the Coxeter number of E_8. So what did their *squares* add to? 120, which he recognised as the cardinality of one of the exceptional finite subgroups of $\mathrm{SU}_2(\mathbb{C})$, and that got him thinking...

Another famous example of A-D-E, due to Arnol'd, are the 'simple critical points' of smooth complex-valued functions f, on e.g. $\mathbb{C}^3$. For example, both $x^2 + y^2 + z^{n+1}$ and $x^2 + y^3 + z^5$ have singularities at $(0, 0, 0)$ (i.e. their first partial derivatives all vanish there), and they are assigned to A_n and E_8, respectively. The $\mathrm{SU}_2(\mathbb{C})$ subgroups can be related to singularities as follows. The group $\mathrm{SU}_2(\mathbb{C})$ acts on $\mathbb{C}^2$ in the obvious way (matrix multiplication). If G is a discrete subgroup, then consider the (ring of) polynomials in 2 variables w_1, w_2 invariant under G. It turns out it will have 3 generators $x(w_1, w_2)$, $y(w_1, w_2)$, $z(w_1, w_2)$, which are connected by 1 polynomial relation (syzygy). For instance, take G to be the cyclic group $\mathbb{Z}_n$, then we're interested in polynomials $p(w_1, w_2)$ invariant under $w_1 \mapsto \exp[2\pi\mathrm{i}/n]w_1$, $w_2 \mapsto \exp[-2\pi\mathrm{i}/n]w_2$. Any such invariant $p(w_1, w_2)$ is clearly generated by (i.e. can be written as a polynomial in) $w_1 w_2$, w_1^n and w_2^n. Choosing instead the generators $x = \frac{w_1^n - w_2^n}{2}$, $y = \mathrm{i}\,\frac{w_1^n + w_2^n}{2}$, $z = w_1 w_2$, we get the syzygy $z^n = -(x^2 + y^2)$. For any G, generators x, y, z can always be found so that the syzygy will be one of the polynomials associated to a simple singularity, and in fact will give the equation of the algebraic surface $\mathbb{C}^2/G$ as a 2-dimensional complex surface in $\mathbb{C}^3$ (e.g. the complex surfaces $\mathbb{C}^2/\mathbb{Z}_n$ and $\{(x, y, z) \in \mathbb{C}^3 \,|\, x^2 + y^2 + z^n = 0\}$ are equivalent).

Arguably the first A-D-E classification goes back to Theaetetus, around 400 B.C. He classified the regular solids. For instance the tetrahedron can be associated to E_6 while the cube is matched with E_7. This A-D-E is only partial, as there are no regular solids assigned to the A-series, and to get the D-series one must look at 'degenerate regular solids'.

The closest thing to an explanation of the A-D-E meta-pattern would seem to be the notion of 'additive assignments' on graphs (which is a picturesque way of describing the corresponding eigenvalue problem). Consider any graph $\mathcal{G}$ with undirected edges, and none of the edges run from a node to itself. We can also assume without loss of generality that $\mathcal{G}$ is connected. Assign a positive number a_i to each node. If this assignment has the property that for each i, $2a_i = \sum a_j$ where the sum is over all nodes j adjacent to i (counting multiplicities of edges), then we call it 'additive'. For instance, for the graph ○═○, the assignment $a_1 = 1 = a_2$ is additive, but the assignment $a_1 = 1, a_2 = 2$ is not. The question is, which graphs have an additive assignment? The answer is: precisely the extended Dynkin diagrams of A-D-E type! And their additive assignments are unique (up to constant proportionality) and are given by the *marks* a_i of the algebra (see e.g. the Table on p.54 of [38]). For example the extended A_n graph consists of $n+1$ nodes arranged in a circle, and its marks a_i all equal 1.

What do additive assignments have to do with the other A-D-E classifications? Consider a finite subgroup G of $\mathrm{SU}_2(\mathbb{C})$. Take the dimension of the equation $G \otimes R_i = \oplus_j m_{ij} R_j$: we get $2d_i = \sum_j m_{ij} d_j$ where $d_j = \dim(R_j)$. Hence the dimensions of the irreducible representations define an additive assignment for each of McKay's graphs, and hence those graphs must be of A-D-E type (provided we know $m_{ij} = m_{ji}$).

As Cappelli-Itzykson-Zuber observed, the physical invariants for $A_1^{(1)}$ also realise the A-D-E pattern, in the following sense. The *Coxeter number* h of the name $\mathcal{X}_\ell$ (i.e. the sum $\sum_i a_i$ of the marks) equals $k+2$, and the *exponents* m_i of X_ℓ equal those $a \in P_+^k$ for which $M_{aa} \neq 0$ (for the simply-laced algebras, the m_i are defined by writing the eigenvalues of the corresponding Cartan matrix (see §2.7) as $4\sin^2(\frac{\pi m_i}{2h})$ — the m_i are integers and the smallest is always 1). Probably what first led Kac to his observation about the E_6 exponents was that $k+2$ (this is how k enters most formulas) for his exceptional equalled the Coxeter number 12 for E_6. More recently, the operator algebraists Ocneanu [48] and independently Böckenhauer-Evans [4] found an A-D-E interpretation for the off-diagonal entries M_{ab} of the $A_1^{(1)}$ physical invariants, using subfactor theory.

We are not claiming that this $A_1^{(1)}$ classification is 'equivalent' to any other A-D-E one — that would miss the point of meta-patterns. What we really want to do is to identify some critical combinatorial part of an $A_1^{(1)}$ proof with critical parts in other A-D-E classifications — this is what we did with the other meta-patterns. A considerably simplified proof of the $A_1^{(1)}$ classification is now available [29], so hopefully this task will now be easier.

There has been some progress at understanding this $A_1^{(1)}$ A-D-E. Nahm [46] constructed the invariant $\mathcal{X}_\ell$ in terms of the compact simply-connected Lie group of type X_ℓ, and in this way could interpret the $k+2=h$ and $M_{m_i m_i} \neq 0$ coincidences. A very general explanation for A-D-E has been suggested by Ocneanu [48] using his theory of path algebras on graphs; although his work has never been published, others are now rediscovering (and publishing!) similar work (see e.g. [4]). Nevertheless, the A-D-E in CFT remains almost as mysterious now as it did a dozen years ago — for example it still isn't clear how it directly relates to additive assignments.

There are 4 other claims for A-D-E classifications of families of RCFT physical invariants, and all of them inherit their (approximate) A-D-E pattern from the more fundamental $A_1^{(1)}$ one. One is the $c<1$ minimal models, also proven in [8], and another is the $N=1$ superconformal minimal models, proved by Cappelli (1987). In both cases the physical invariants are parametrised by pairs of A-D-E diagrams. The list of known $c=1$ RCFTs also looks like A-D-E (two series parametrised by $\mathbb{Q}_+$, and three exceptionals), but the completeness of that list has never been rigourously established.

The fourth classification often quoted as A-D-E, is the $N=2$ superconformal minimal models. Their classification was done by Gannon (1997). The connection here with A-D-E turns out to be rather weak: e.g. 20, 30, and 24 distinct invariants would have an equal right to be called $\mathcal{E}_6$, $\mathcal{E}_7$, and $\mathcal{E}_8$ respectively. It would appear that the frequent claims that the $N=2$ minimal models fall into an A-D-E pattern are rather dubious.

Hanany-He [35] suggest that the $A_1^{(1)}$ A-D-E pattern can be related to subgroups $G \subset \mathrm{SU}_2(\mathbb{C})$ by orbifolding 4-dimensional $N=4$ supersymmetric gauge theory by G, resulting in an $N=2$ superCFT whose 'matter matrix' can be read off from the Dynkin diagram corresponding to G. The same game can be played with finite subgroups of $\mathrm{SU}_3(\mathbb{C})$, resulting in $N=1$ superCFTs whose matter matrices correspond to graphs very reminiscent of the 'fusion graphs' of Di Francesco-Petkova-Zuber (see e.g. [62]) corresponding to $A_2^{(1)}$ physical invariants. [35] use this to conjecture a McKay-type correspondence between singularities of type $\mathbb{C}^n/G$, for $G \subset \mathrm{SU}_n(\mathbb{C})$, and the physical invariants of $A_{n-1}^{(1)}$. This in their view would be the form A-D-E takes for higher rank physical invariants. Their actual conjecture though is still somewhat too vague.

For a final example of meta-pattern, consider 'modular function' (see §2.3). After all, they appear in a surprising variety of places and disguises. Maybe we shouldn't regard their ubiquity as fortuitous, instead perhaps there's a deeper common 'situation' which is the source for that ubiquity. Just as 'symmetry' yields 'group', or 'rain-followed-by-heat' breeds mosquitos. Math is not above metaphysics; like any area it grows by asking questions, and changing your perspective — even to a metaphysical one — should suggest new questions.

1.7. SIMPLE-CURRENTS AND CHARGE-CONJUGATION

The key properties[15] of the matrix S are that it's unitary and symmetric (so M in §1.5 equals SMS^*),

$$S_{0\mu} > 0 \text{ for all } \mu \in P_+^k \ , \tag{1.7.1}$$

and that the numbers $N_{\lambda\mu}^{\nu}$ defined by Verlinde's formula (1.4.4) are nonnegative integers. These are obeyed by the matrix S in any (unitary) RCFT. From these basic properties, we will obtain here some elementary consequences which have important applications.

But first, let's make an observation which isn't difficult to prove, but doesn't appear to be generally known.

Verlinde's formula looks strange, but it is quite generic,

and we can see it throughout math and mathematical physics. Consider the following.

Let $\mathfrak{A}$ be a commutative associative algebra, over $\mathbb{R}$ say. Suppose $\mathfrak{A}$ has a finite basis Φ (over $\mathbb{R}$) containing the unit 1. Define the 'structure constants' $N_{ab}^c \in \mathbb{R}$, for $a,b,c \in \Phi$, by $ab = \sum_{c\in\Phi} N_{ab}^c c$. Suppose there is an algebra homomorphism $*$ (so $*$ is linear, and $(xy)^* = x^*y^*$) which permutes the basis vectors (so $\Phi^* = \Phi$), and we have the relation $N_{ab}^1 = \delta_{b,a^*}$. We call any such algebra $\mathfrak{A}$ a *fusion algebra.* Then any fusion algebra will necessarily have a unitary matrix S with $S_{1a} > 0$ and with the structure constants given by Verlinde's formula. Algebraically, the relation $S = S^t$ holds if $\mathfrak{A}$ is 'self-dual' in a certain natural sense.

---

[15] A good exercise for the reader is to prove that if S is unitary and symmetric, and obeys (1.7.1), then there will be at most finitely many physical invariants M for that S,T.

Define the 'fusion matrices' N_λ by $(N_\lambda)_{\mu\nu} = N_{\lambda\mu}^{\nu}$. Then Verlinde's formula says that the μth column $S_{\uparrow\mu}$ of S is an eigenvector of each fusion matrix N_λ, with eigenvalue $\frac{S_{\lambda\mu}}{S_{0\mu}}$.

Useful Fact. If v is a simultaneous eigenvector of each fusion matrix N_λ, then there exists a constant $c \in \mathbb{C}$ and a $\lambda \in P_+^k$ such that $v = c\, S_{\uparrow\lambda}$.

For one consequence, take the complex conjugate of the eigenvector equation $N_\lambda\, S_{\uparrow\mu} = \frac{S_{\lambda\mu}}{S_{0\mu}}\, S_{\uparrow\mu}$: we get that the vector $S_{\uparrow\mu}^*$ is a simultaneous eigenvector of all N_λ, and hence must equal $c\, S_{\uparrow\gamma}$ for some number c and weight $\gamma \in P_+^k$, both depending on μ. Write $\gamma = C\mu$; then C defines a permutation of P_+^k. The reader can verify that unitarity of S forces $|c| = 1$, while (1.7.1) forces $c > 0$. Thus $c = 1$ and we obtain the formula

$$S_{\lambda\mu}^* = S_{\lambda,C\mu} = S_{C\lambda,\mu}\ . \tag{1.7.2}$$

Also, unitarity and symmetry of S forces $C = S^2$, while conjugating twice shows $C^2 = id$. C is an important matrix in RCFT, and is called *charge-conjugation*. When $C = id.$, then the matrix S is real.

Note that (1.7.1) now implies $C0 = 0$. Also $CT = TC$. Hence $M = C$ will always define a physical invariant, and if M is any other physical invariant, the matrix product $MC = CM$ will define another physical invariant. Also, $N_{C\lambda,C\mu}^{C\nu} = N_{\lambda\mu}^{\nu}$ and $(N_\lambda)^t = CN_\lambda = N_\lambda C = N_{C\lambda}$.

For the WZW (=affine) case, C has a special meaning: $C\lambda$ is the highest-weight 'contragredient' to λ. C corresponds to an order 2 (or 1) symmetry of the (unextended) Dynkin diagram. For example, for $A_\ell^{(1)}$, we have $C(\lambda_0, \lambda_1, \ldots, \lambda_{\ell-1}, \lambda_\ell) = (\lambda_0, \lambda_\ell, \lambda_{\ell-1}, \ldots, \lambda_1)$. For $A_1^{(1)}$ then, $C = id.$, which can also be read off from (1.4.3).

The algebras $D_{even}^{(1)}$ all have at least one nontrivial symmetry of the (unextended) Dynkin diagram which isn't the charge-conjugation. The most interesting example is $D_4^{(1)}$, which has 5 of these. By a *conjugation*, we will mean any symmetry of the unextended Dynkin diagram.

To go much further, we need a fascinating tool called *Perron-Frobenius theory* — a collection of results concerning the eigenvalues and eigenvectors of nonnegative matrices (i.e. matrices in which every entry is a nonnegative real number). Whenever you have such matrices in your problem, and it is natural to multiply them, then there is a good chance Perron-Frobenius theory will tell you something interesting. The basic result here is that if A is a nonnegative matrix, then there will be a nonnegative eigenvector $x \geq 0$ with eigenvalue $\rho \geq 0$, such that if λ is any other eigenvalue of A, then $|\lambda| \leq \rho$. There are lots of other results (see e.g. [44]), e.g. ρ must be at least as large as any diagonal entry of A, and there must be a row-sum of A no bigger than ρ, and another row-sum no smaller than ρ.

For instance, consider

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\ , \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\ .$$

Perron-Frobenius eigenvectors for A and B are $\begin{pmatrix}1\\1\\1\end{pmatrix}$ and $\begin{pmatrix}2\\1\\1\end{pmatrix}$, with eigenvalues 3 and 2 resp. The other eigenvalue of A is 0 (multiplicity 2), while those of B are 0 and -1.

Fusion matrices N_λ are nonnegative, and it is indeed natural to multiply them:

$$N_\lambda N_\mu = \sum_{\nu\in P_+^k} N_{\lambda\mu}^\nu N_\nu \ .$$

So we can expect Perron-Frobenius to tell us something interesting. This is the case, and we obtain the curious-looking inequalities

$$S_{\lambda 0}\, S_{0\mu} \geq |S_{\lambda\mu}|\, S_{00} \ . \tag{1.7.3}$$

Squaring both sides, summing over μ and using unitarity, we get that $S_{\lambda 0} \geq S_{00}$. In other words, the ratio $\frac{S_{\lambda 0}}{S_{00}}$, called the *quantum-dimension* of λ, will necessarily be ≥ 1.

The term 'quantum-dimension' comes from quantum groups, where $\frac{S_{\lambda 0}}{S_{00}}$ is the quantum-dimension of the module labelled by λ of the quantum group $U_q(X_\ell)$.

The borderline case then [16] is when a quantum-dimension *equals* 1. Any such weight is called a *simple-current*. The theory of simple-currents was developed most extensively by Schellekens and collaborators (see e.g. [50]). The simple-currents for the affine algebras were classified by J. Fuchs (1991), and the result is that (with one unimportant exception: $E_8^{(1)}$ at level 2) they all correspond to symmetries of the extended Dynkin diagrams. In particular, applying any such symmetry to the vacuum $0 = (k, 0, \ldots, 0)$ gives the list of simple-currents. For instance, the $\ell + 1$ weights of the form $(0, \ldots, k, \ldots, 0)$ (k in the ith spot) are the simple-currents for $A_\ell^{(1)}$. There are 2 simple-currents for $B_\ell^{(1)}$, $C_\ell^{(1)}$ and $E_7^{(1)}$, 3 for $E_6^{(1)}$, and 4 for $D_\ell^{(1)}$. Simple-currents play a large role in RCFT, as we shall see.

Let j be any simple-current. Then (1.7.3) becomes $S_{0\mu} \geq |S_{j\mu}|$ for all μ, so unitarity forces $S_{0\mu} = |S_{j\mu}|$, that is

$$S_{j\mu} = \exp[2\pi \mathrm{i}\, Q_j(\mu)]\, S_{0\mu} \qquad \forall \mu \in P_+^k \tag{1.7.4}$$

for some rational numbers $0 \leq Q_j(\mu) < 1$. Hence by diagonalising, we get $N_j N_{Cj} = I$. But the inverse of a nonnegative matrix A is itself nonnegative, only if A is a 'generalised permutation matrix', i.e. a permutation matrix except the 1's can be replaced by any positive numbers. But N_j and N_{Cj} are also integral, and so they must in fact be permutation matrices. Write $(N_j)_{\lambda\mu} = \delta_{\mu,J\lambda}$ for some permutation J of P_+^k. So $j = J0$. Then

$$\delta_{\lambda,\mu} = N_{j,\lambda}^{J\mu} = \sum_\nu \exp[2\pi \mathrm{i}\, Q_j(\nu)]\, S_{\lambda\nu}\, S_{J\mu,\nu}^* \ ,$$

---

[16] This seems to be a standard trick in math: when some sort of bound is established, look at the extremal cases which realise that bound. If your bound is a good one, it should be possible to say something about those extremal cases, and having something to say is always of paramount importance. This trick is used for instance in the definition of 24 last section, and the definition of normal subgroup in section 2.2.

so taking absolute values and using the triangle inequality and unitarity of S, we find that (1.7.4) generalises:

$$S_{J\lambda,\mu} = \exp[2\pi\mathrm{i}\, Q_j(\mu)]\, S_{\lambda\mu}\ . \tag{1.7.5}$$

The simple-currents form a finite abelian group, corresponding to the composition of the permutations J. For any simple-currents J, J', we get the symmetry $N_{J\lambda,J'\mu}^{\nu} = N_{\lambda\mu}^{JJ'\nu}$. The $\mathbb{Q}/\mathbb{Z}$-valued functions Q_j define gradings on the fusion rings, and conversely any grading corresponds to a simple-current in this way.

For example, the simple-current $j = (0, k)$ of $A_1^{(1)}$ at level k corresponds to $Q_j(\lambda) = \lambda_1/2$ and the permutation $J\lambda = (\lambda_1, \lambda_0)$. We can see this directly from (1.4.3). For $A_2^{(1)}$ level k, there are 2 nontrivial simple-currents, $(0, k, 0)$ and $(0, 0, k)$. The first of these corresponds to triality $Q(\lambda) = (\lambda_1 + 2\lambda_2)/3$ and $\lambda \mapsto (\lambda_2, \lambda_0, \lambda_1)$, while the second to $(2\lambda_1 + \lambda_2)/3$ and $\lambda \mapsto (\lambda_1, \lambda_2, \lambda_0)$. Similar statements hold for all affine algebras: e.g. for $B_\ell^{(1)}$ level k, the nontrivial simple-current has $Q_j(\lambda) = \lambda_\ell/2$ and $J\lambda = (\lambda_1, \lambda_0, \lambda_2, \ldots, \lambda_\ell)$.

One of the applications of simple-currents is that physical invariants can be built from them in generic ways. These physical invariants all obey the selection rule

$$M_{\lambda\mu} \neq 0 \quad \Longrightarrow \quad \mu = J\lambda \quad \text{for some simple}-\text{current } J = J(\lambda, \mu)\ . \tag{1.7.6}$$

We will call any such physical invariant M a *simple-current invariant.* A special case is the $\mathcal{D}_{\frac{k}{2}+2}$ physical invariant for $A_1^{(1)}$ at even level k. Up to a fairly mild assumption, all simple-current invariants have been classified for any RCFT by Schellekens and collaborators; given that assumption, they can all be constructed by generic methods. The basic construction is due to Bernard [3], though it has been generalised by others. In the WZW case, all simple-current invariants (except some for $D_\ell^{(1)}$) correspond to strings on nonsimply-connected Lie groups.

By a *generic physical invariant* of $X_\ell^{(1)}$ we mean one of the form $M = C'M'C''$ where C', C'' are (charge-)conjugations, and M' is a simple-current invariant. In other words, M is constructed in generic ways from symmetries of the extended Dynkin diagram of X_ℓ. Any other M are called *exceptional.*

All known results point to the validity of the following guess:

Conjecture. Choose a simple algebra X_ℓ. Then for all sufficiently large k, all physical invariants of $X_\ell^{(1)}$ at level k will be generic.

In other words, any given $X_\ell^{(1)}$ will have only finitely many exceptionals. For instance, for $A_1^{(1)}$ and $A_2^{(1)}$ at any $k > 28$ and $k > 21$ resp., all physical invariants are generic. For $C_2^{(1)}$ and $G_2^{(1)}$, $k > 12$ and $k > 4$ resp. should work.

The richest source of exceptionals are *conformal embeddings.* In some cases the affine representations L_λ for some algebra $X_\ell^{(1)}$ (necessarily at level 1) can be decomposed into *finite* direct sums of representations of some affine subalgebra $Y_m^{(1)}$ (at some level k). In this case, a physical invariant for $X_\ell^{(1)}$ level 1 will yield a physical invariant for $Y_m^{(1)}$ level k, obtained by replacing every $X_\ell^{(1)}$ level 1 character χ_λ by the appropriate finite sum of

$Y_m^{(1)}$ level k characters. An example will demonstrate this simple idea: $A_1^{(1)}$ level 28 is a conformal subalgebra of $G_2^{(1)}$ level 1, and we have the character decompositions

$$\begin{aligned}\chi_{(1,0,0)} &= \chi'_{(28,0)} + \chi'_{(18,10)} + \chi'_{(10,18)} + \chi'_{(0,28)} \\ \chi_{(0,0,1)} &= \chi'_{(22,6)} + \chi'_{(16,12)} + \chi'_{(12,16)} + \chi'_{(6,22)} \ .\end{aligned}$$

Thus the unique level 1 $G_2^{(1)}$ physical invariant $|\chi_0|^2 + |\chi_{(0,0,1)}|^2$ yields what we call the $\mathcal{E}_8$ physical invariant of $A_1^{(1)}$. All level 1 physical invariants are known, as are all conformal embeddings and the corresponding character decompositions (branching rules).

1.8. Galois Theory

Evariste Galois was a brilliantly original French mathematician. Born shortly before Napoleon's ill-fated invasion of Russia, he died shortly before the ill-fated 1832 uprising in Paris. His last words: "Don't cry, I need all my courage to die at 20".

Galois grew up in a time and place confused and excited by revolution. He was known to say "if I were only sure that a body would be enough to incite the people to revolt, I would offer mine". On May 2 1832, after frustration over failure in love and failure to convince the Paris math establishment of the depth of his ideas, he made his decision. A duel was arranged with a friend, but only his friend's gun would be loaded. Galois died the day after a bullet perforated his intestine. At his funeral it was discovered that a famous general had also just died, and the revolutionaries decided to use the general's death rather than Galois' as a pretext for an armed uprising. A few days later the streets of Paris were blocked by barricades, but not because of Galois' sacrifice: his death had been pointless [56].

Galois theory in its most general form is the study of relations between objects defined implicitly by some conditions. For example, the objects could be the solutions to a given differential equation. In the incarnation of Galois we are interested in here, the objects are numbers, namely the zeros of certain polynomials. We will sketch this theory below, but see e.g. the article by Stark in [60] for more details.

Gauss seems to have been the first to show that 'weird' (complex) numbers could tell us about the integers. For instance, suppose we are interested in the equation $n = a^2 + b^2$. Consider $5 = 2^2 + 1^2$. We can write this as $5 = (2 + \mathrm{i})(2 - \mathrm{i})$, so we are led to consider complex numbers of the form $a + b\mathrm{i}$, for $a, b \in \mathbb{Z}$. These are now called 'Gaussian integers'. Suppose we know the following theorem:

Fact. Let $p \in \mathbb{Z}$ be any prime number. Then p factorises over the Gaussian integers iff $p = 2$ or $p \equiv 1 \pmod 4$.

By 'factorise' there, we mean $p = zw$ where neither z nor w is a 'unit': ± 1, $\pm \mathrm{i}$.

Now suppose p is a prime, $= 2$ or $\equiv 1 \pmod 4$, and we write $p = (a + b\mathrm{i})(c + d\mathrm{i})$. Then $p^2 = (a^2 + b^2)(c^2 + d^2)$, so $a^2 + b^2 = c^2 + d^2 = p$. Conversely, suppose $p = a^2 + b^2$, then $p = (a + b\mathrm{i})(a - b\mathrm{i})$. Thus:

Consequence. [17] Let $p \in \mathbb{Z}$ be any prime number. Then $p = a^2 + b^2$ for $a, b \in \mathbb{Z}$ iff $p = 2$ or $p \equiv 1 \pmod 4$.

Now we can answer the question: can a given n be written as a sum of 2 squares $n = a^2 + b^2$? Write out the prime decomposition $n = \prod p^{a_p}$. Then $n = a^2 + b^2$ has a solution iff a_p is *even* for every $p \equiv 3 \pmod 4$. For instance $60 = 2^2 \cdot 3^1 \cdot 5^1$ cannot be written as the sum of 2 squares, but $90 = 2^1 \cdot 3^2 \cdot 5^1$ can. We can also find (and count) all solutions: e.g. $90 = 2 \cdot 3^2 \cdot 5 = \{(1+\mathrm{i})3(1+2\mathrm{i})\}\{(1-\mathrm{i})3(1-2\mathrm{i})\}$, giving $90 = (-3)^2 + 9^2$.

This problem should give the reader a small appreciation for the power of using nonintegers to study integers. Nonintegers often lurk in the shadows, secretly watching their more arrogant brethren the integers strut. One of the consequences of their presence can be the existence of certain 'Galois' symmetries. Such happens in RCFT, as we will show below.

Look at complex conjugation: $(wz)^* = w^* z^*$ and $(w + z)^* = w^* + z^*$. Also, $r^* = r$ for any $r \in \mathbb{R}$. So we can say that $*$ is a structure-preserving map $\mathbb{C} \to \mathbb{C}$ (called an *automorphism* of $\mathbb{C}$) fixing the reals. We will write this $* \in \mathrm{Gal}(\mathbb{C}/\mathbb{R})$. '$\mathrm{Gal}(\mathbb{C}/\mathbb{R})$' is the Galois group of $\mathbb{C}$ over $\mathbb{R}$; it turns out to contain only $*$ and the identity.

A way of thinking about the automorphism $*$ is that it says that, as far as the real numbers are concerned, i and $-$i are identical twins.

Let $\mathbb{F}$ be any *field* containing $\mathbb{Q}$ (we defined 'field' in §1.2). The *Galois group* $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ then will be the set of all automorphisms=symmetries of $\mathbb{F}$ which fix all rationals.

For example, take $\mathbb{F}$ to be the set of all numbers of the form $a + b\sqrt{5}$, where $a, b \in \mathbb{Q}$. Then $\mathbb{F}$ will be a field, which is commonly denoted $\mathbb{Q}[\sqrt{5}]$ because it is generated by $\mathbb{Q}$ and $\sqrt{5}$. Let's try to find its Galois group. Let $\sigma \in \mathrm{Gal}(\mathbb{F}/\mathbb{Q})$. Then $\sigma(a + b\sqrt{5}) = \sigma(a) + \sigma(b)\sigma(\sqrt{5}) = a + b\sigma(\sqrt{5})$, so once we know what σ does to $\sqrt{5}$, we know everything about σ. But $5 = \sigma(5) = \sigma(\sqrt{5}^2) = (\sigma(\sqrt{5}))^2$, so $\sigma(\sqrt{5}) = \pm\sqrt{5}$ and there are precisely 2 possible Galois automorphisms here (one is the identity). As far as $\mathbb{Q}$ is concerned, $\pm\sqrt{5}$ are interchangeable: it cannot see the difference.

For a more important example, consider the *cyclotomic field* $\mathbb{F} = \mathbb{Q}[\xi_n]$, where $\xi_n := \exp[2\pi\mathrm{i}/n]$ is an nth root of 1. So $\mathbb{Q}[\xi_n]$ consists of all complex numbers which can be expressed as polynomials $a_m\xi_n^m + a_{m-1}\xi_n^{m-1} + \cdots + a_0$ in ξ_n with rational coefficients a_i. Once again, to find the Galois group $\mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$, it is enough to see what an automorphism σ does to the generator ξ_n. Since $\xi_n^n = 1$, we see that it must send it to another nth root of 1, ξ_n^ℓ say; in fact it is easy to see that $\sigma(\xi_n)$ must be another 'primitive' nth root of 1, i.e. ℓ must be coprime to n. So $\mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$ will be isomorphic to the multiplicative group $\mathbb{Z}_n^\times$ of numbers between 1 and n coprime to n. The rationals can't see any difference between the primitive nth roots of 1 — for instance $\mathbb{Q}$ can't tell that $\xi_n^{\pm 1}$ are 'closer to 1' than the other primitive roots. So any $\sigma \in \mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$ will correspond to some $\ell \in \mathbb{Z}_n^\times$, and to see what σ does to some $z \in \mathbb{Q}[\xi_n]$ what we do is write z as a polynomial in ξ_n and then replace each occurrence of ξ_n with ξ_n^ℓ. For example,

$$\sigma\big(\cos(2\pi a/n)\big) = \sigma\left(\frac{\xi_n^a + \xi_n^{-a}}{2}\right) = \frac{\xi_n^{a\ell} + \xi_n^{-a\ell}}{2} = \cos(2\pi a\ell/n) \ .$$

---

[17] This result was first stated by Fermat in one of his infamous margin notes (another is discussed in Section 2.3), and was finally proved a century later by Euler. A remarkable 1-line proof was found by Zagier [61].

So to summarise, Galois automorphisms are a massive generalisation of the idea of complex conjugation. If in your problem complex conjugation seems interesting, then there is a good chance more general Galois automorphisms will play an interesting role. This is what happens in RCFT, as we now show.

Fact. [14] Suppose S is unitary and symmetric and each $S_{0a} > 0$.

(a) If in addition the numbers N_{ab}^c given by Verlinde's formula (1.4.4) are rational, then the entries S_{ab} of S must lie in a cyclotomic field.

(b) The numbers N_{ab}^c will be rational iff for any $\sigma \in \mathrm{Gal}(\mathbb{Q}[S]/\mathbb{Q})$, there is a permutation $a \mapsto a^\sigma$, and a choice of signs $\epsilon_\sigma(a) \in \{\pm 1\}$, such that

$$\sigma(S_{ab}) = \epsilon_\sigma(a)\, S_{a^\sigma,b} = \epsilon_\sigma(b)\, S_{a,b^\sigma} \ . \tag{1.8.1}$$

'$\mathbb{Q}[S]$' in part (b) denotes the field generated by $\mathbb{Q}$ and all matrix entries S_{ab}. The argument follows the one given for the charge-conjugation C at the beginning of the last section. The kinds of complex numbers which lie in cyclotomic fields are $\sin(\pi r)$, $\cos(\pi r)$, $\sqrt{r}$ and $r\mathrm{i}$ for any $r \in \mathbb{Q}$. Almost all complex numbers fail to lie in any cyclotomic field: e.g. generic cube roots, 4th roots, ..., of rationals, as well as transcendental numbers like e, π and e^π.

Of course the affine algebras satisfy the conditions of the Fact, as does more generally the modular matrix S for any unitary RCFT, and so these will possess the Galois action. For the affine algebras this action has a geometric interpretation in terms of multiplying weights by an integer ℓ and applying Weyl group elements — see [14] for a description.

This Fact is useful in both directions: as a way of testing whether a conjectured matrix S has a chance of producing the integral fusions we want it to yield; and more importantly as a source of a symmetry of the RCFT which generalises charge-conjugation. Any statement about charge-conjugation seems to have an analogue for any of these Galois symmetries, although it is usually more complicated.

As an example, consider $A_1^{(1)}$: (1.4.3) shows explicitly that $S_{\lambda\mu}$ lies in the cyclotomic field $\mathbb{Q}[\xi_{4(k+2)}]$. Write $\{x\}$ for the number congruent to x mod $2(k+2)$ satisfying $0 \le \{x\} < 2(k+2)$. Choose any Galois automorphism σ, and let $\ell \in \mathbb{Z}_{4(k+2)}^\times$ be the corresponding integer. Then if $\{\ell(a+1)\} < k+2$, we will have $a^\sigma = \{\ell(a+1)\}-1$, while if $\{\ell(a+1)\} > k+2$, we'll have $a^\sigma = 2(k+2) - \{\ell(a+1)\} - 1$. The sign $\epsilon_\sigma(a)$ will depend on a contribution from $\sqrt{\frac{2}{k+2}}$ (which for most purposes can be ignored), as well as the sign $+1$ or -1, resp., depending on whether or not $\{\ell(a+1)\} < k+2$.

Consider specifically $k = 10$, and the Galois automorphism σ_5 corresponding to $\ell = 5$. Then the permutation is $0 \leftrightarrow (6,4)$, $(9,1) \leftrightarrow (1,9)$, $(8,2) \leftrightarrow (2,8)$, $(4,6) \leftrightarrow (0,10)$, while $(7,3)$ and (5,5) are fixed.

This Galois symmetry has been used to find certain exceptional physical invariants, but its greatest use so far is as a powerful selection rule we will describe next section.

In this final section we include some of the basic tools belonging to the 'modern' classifications of physical invariants, and we give a flavour of their proofs. We will state them for the $A_1^{(1)}$ level k problem given above, but everything generalises without effort. See [29] and references therein for more details. Recall the matrices S, T in (1.4.3).

First note that commutation of M with T implies the selection rule

$$M_{\lambda\mu} \neq 0 \quad \Longrightarrow \quad (\lambda_1+1)^2 \equiv (\mu_1+1)^2 \pmod{4(k+2)} \ . \tag{1.9.1}$$

It is much harder to squeeze information out of the commutation with S, but the resulting information turns out surprisingly to be much more useful. In fact, commutation with S is almost incompatible with the constraint $M_{\lambda\mu} \in \mathbb{Z}_{\geq}$.

Note that the vacuum $0 \in P_+^k$ is both physically and mathematically special; our strategy will be to find all possible 0th rows and columns of M, and then for each of these possibilities to find the remaining entries of M.

The easiest result follows by evaluating $MS = SM$ at $(0, \lambda)$ for any $\lambda \in P_+^k$:

$$\sum_{\mu \in P_+^k} M_{0\mu}\, S_{\mu\lambda} \geq 0 \ , \tag{1.9.2}$$

with equality iff the λth column of M is identically 0. (1.9.2) has two uses: it severely constrains the values of $M_{0\mu}$ (similarly $M_{\mu 0}$), and it says precisely which columns (and rows) are nonzero.

Next, let's apply the triangle inequality to sums involving (1.7.5). Choose any $i, j \in \{0, 1\}$. Then

$$M_{J^i 0, J^j 0} = \sum_{\lambda,\mu} (-1)^{\lambda_1 i}\, S_{0\lambda}\, M_{\lambda\mu}\, (-1)^{\mu_1 j}\, S_{0\mu} \ .$$

Taking absolute values, we obtain

$$M_{J^i 0, J^j 0} \leq \sum_{\lambda,\mu} S_{0\lambda}\, M_{\lambda\mu}\, S_{0\mu} = M_{00} = 1 \ .$$

Thus $M_{J^i 0, J^j 0}$ can equal only 0 or 1. If it equals 1, then we obtain the selection rule:

$$\lambda_1 i \equiv \mu_1 j \pmod 2 \text{ whenever } M_{\lambda\mu} \neq 0 \ ;$$

this implies the symmetry $M_{J^i\lambda, J^j\mu} = M_{\lambda\mu}$ for all $\lambda, \mu \in P_+^k$. We can see both of these in the list of physical invariants for $A_1^{(1)}$ level k. This explains a lot of the properties of those invariants. For instance, try to use this selection rule to explain why no χ_{odd} appears in the exceptional called $\mathcal{E}_8$.

Our M is nonnegative, and although multiplying M's may not give us back a physical invariant, it will give us a matrix commuting with S and T. In other words, the commutant is much more than merely a vector space, it is in fact an algebra. Thus we should expect Perron-Frobenius to tell us something here. A first application is the following.

Suppose $M_{\lambda 0}=\delta_{\lambda,0}$ — i.e. the 0th column of M is all zeros except for $M_{00}=1$. Then Perron-Frobenius implies (with a little work) that M will be a permutation matrix — i.e. there is some permutation π of P_+^k such that $M_{\lambda\mu}=\delta_{\mu,\pi\lambda}$, and $S_{\pi\lambda,\pi\mu}=S_{\lambda\mu}$. This nice fact applies directly to the $\mathcal{A}_\star$ and $\mathcal{D}_{odd}$ physical invariants of $A_1^{(1)}$.

This is proved by studying the powers $(M^t\,M)^L$ as L goes to infinity: its diagonal entries will grow exponentially with L, unless there is at most one nonzero entry on each row of M, and that entry equals 1.

More careful reasoning along those lines tells us about the other generic situation here. Namely, suppose $M_{\lambda 0}\neq 0$ only for $\lambda=0$ and $\lambda=J0$, and similarly for $M_{0\lambda}$ — i.e. the 0th row and column of M are all zeros except for $M_{J^i0,J^j0}=1$. Then the λth row (or column) of M will be identically 0 iff λ_1 is odd. Moreover, let λ,μ be any non-fixed-points of J, and suppose $M_{\lambda\mu}\neq 0$. Then

$$M_{\lambda\nu}=\begin{cases}1 & \text{if } \nu=\mu \text{ or } \nu=J\mu\\ 0 & \text{otherwise}\end{cases}$$

with a similar formula for $M_{\nu\mu}$. This applies to the $\mathcal{D}_{even}$ and $\mathcal{E}_7$ invariants of $A_1^{(1)}$.

Our final ingredient is the Galois symmetry (1.8.1) obeyed by S. Choose any Galois automorphism σ. It will correspond to some integer ℓ coprime to $2(k+2)$. From (1.8.1) and $M=SMS^*$ we get, for all λ,μ, the important relation

$$M_{\lambda\mu}=\epsilon_\sigma(\lambda)\,\epsilon_\sigma(\mu)\,M_{\lambda^\sigma,\mu^\sigma}\ . \tag{1.9.3}$$

From (1.9.3) and the positivity of M, we obtain the powerful *Galois selection rule*

$$M_{\lambda\mu}\neq 0 \quad\Longrightarrow\quad \epsilon_\sigma(\lambda)=\epsilon_\sigma(\mu)\ . \tag{1.9.4}$$

Next let us quickly sketch how these tools are used to obtain the $A_1^{(1)}$ classification. For details the reader should consult [29].

The first step will be to find all possible values of λ such that $M_{0\lambda}\neq 0$ or $M_{\lambda 0}\neq 0$. These λ are severely constrained. We know two generic possibilities: $\lambda_1=0$ (good for all k), and $\lambda=J0$ (good when $\frac{k}{2}$ is even). We now ask the question, what other possibilities for λ are there? Our goal is to prove (1.9.7). Assume $\lambda\neq 0,J0$, and write $a=\lambda_1+1$ and $n=k+2$.

There are only two constraints on λ which we will need. One is (1.9.1):

$$(a-1)\,(a+1)\equiv 0 \pmod{4n}\ . \tag{1.9.5}$$

More useful is the Galois selection rule (1.9.4), which we can write as $\sin(\pi\ell\frac{a}{n})\sin(\pi\ell\frac{1}{n})>0$, for all those ℓ. But a product of sines can be rewritten as a difference of cosines, so

$$\cos(\pi\,\ell\,\frac{a-1}{n})>\cos(\pi\,\ell\,\frac{a+1}{n})\ . \tag{1.9.6}$$

(1.9.6) is strong and easy to solve; the reader should try to find her own argument.

What we get is that, provided $n \neq 12, 30$, M obeys the strong condition

$$M_{\lambda 0} \neq 0 \text{ or } M_{0\lambda} \neq 0 \qquad \Longrightarrow \qquad \lambda \in \{0, J0\} \ . \tag{1.9.7}$$

Consider first **case 1**: $M_{\lambda 0} = \delta_{\lambda,0}$. From above, we know $M_{\lambda\mu} = \delta_{\mu,\pi\lambda}$ for some permutation π of P_+^k obeying $S_{\lambda\mu} = S_{\pi\lambda,\pi\mu}$. We know $\pi 0 = 0$; put $\mu := \pi(k-1,1)$. Then $\sin(\pi\frac{2}{n}) = \sin(\pi\frac{\mu_1+1}{n})$, and so we get either $\mu = (k-1,1)$ or $\mu = J(k-1,1)$. By T-invariance (1.9.1), the second possibility can only occur if $4 \equiv (n-2)^2 \pmod{4n}$, i.e. 4 divides n. But for those n, $\mathcal{D}_{\frac{n}{2}+1}$ is also a permutation matrix, so replacing M if necessary with the matrix product $M\,\mathcal{D}_{\frac{n}{2}+1}$, we can always require $\mu = (k-1,1)$, i.e. π also fixes $(k-1,1)$. It is now easy to show π must fix any λ, i.e. that M is the identity matrix $\mathcal{A}_{n-1}$.

The other possibility, **case 2**, is that both $M_{0,J0} \neq 0$ and $M_{J0,0} \neq 0$. (1.9.1) says $1 \equiv (n-1)^2 \pmod{4n}$, i.e. $\frac{n}{2}$ is odd. The argument here is similar to that of **case 1**, but with $(k-2,2)$ playing the role of $(k-1,1)$. We can show that $M_{(k-2,2),(k-2,2)} \neq 0$, except possibly for $k = 16$, where we find the exceptional $\mathcal{E}_7$. Otherwise we get $M = \mathcal{D}_{\frac{n}{2}+1}$.

For more general $X_\ell^{(1)}$ level k, the approach is

(i) to look at all the constraints on the $\lambda \in P_+^k$ for which $M_{0\lambda} \neq 0$ or $M_{\lambda 0} \neq 0$. Most important here are $TM = MT$ (which will always be some sort of norm selection rule) and the Galois selection rule (1.9.4). Generically, what we will find is that such a λ must equal $J0$ for some simple-current J, as in (1.9.7) for $A_1^{(1)}$.

(ii) Solve this generic case (in the $A_1^{(1)}$ classification, these were the physical invariants $\mathcal{A}_\star$, $\mathcal{D}_\star$ and $\mathcal{E}_7$).

(iii) Solve the nongeneric case. The worst of these are the orthogonal algebras at $k = 2$, as well as the places where conformal embeddings (see §1.7) occur.

(ii) has recently been completed for all simple X_ℓ, as has the $k = 2$ part of (iii). (i) is the main remaining task in the physical invariant classification for simple X_ℓ.

A natural question to ask is whether A-D-E has been observed in e.g. the $A_2^{(1)}$ classification. The answer is no, although the fusion graph theory of Di Francesco-Petkova-Zuber [62] is an attempt to assign to these physical invariants graphs reminiscent of the A-D-E Dynkin diagrams. Also, there is related work trying to understand the $A_2^{(1)}$ classification in terms of subgroups of $\mathrm{SU}_3(\mathbb{C})$ (as opposed to $\mathrm{SU}_2(\mathbb{C})$ for $A_1^{(1)}$) — see e.g. [35]. Finding the $A_3^{(1)}$, $A_4^{(1)}$,... classifications would permit the clarification and testing of this vaguely conjectured relation between the $A_n^{(1)}$ physical invariants, and singularities $\mathbb{C}^{n+1}/G$ for G a finite subgroup of $\mathrm{SU}_{n+1}(\mathbb{C})$.

However, a few years ago Philippe Ruelle was walking in a library in Dublin. He spotted a yellow book in the math section, called *Complex Multiplication* by Lang. A strange title for a book by Lang! After all, there can't be all that much even Lang could really say about complex multiplication! Ruelle flipped it to a random page, which turned out to be p.26. On there he found what we would call the Galois selection rule for $A_2^{(1)}$, analysed and solved for the cases where $k+3$ is coprime to 6. Lang however didn't know about physical invariants; he was reporting on work by Koblitz and Rohrlich on decomposing the Jacobians of the Fermat curve $x^n + y^n = z^n$ into their prime pieces, called 'simple factors' in algebraic geometry. n here corresponds to $k+3$. Similarly, Itzykson discovered traces of the $A_2^{(1)}$ exceptionals — these occur when $k+3 = 8, 12, 24$ — in the Jacobian of $x^{24} + y^{24} = z^{24}$. See [2] for further observations along these lines. These 'coincidences' are still far from understood. Nor is it known if, more generally, the $A_\ell^{(1)}$ level k classification will somehow be related to the hypersurface $x_1^n + \cdots + x_\ell^n = z^n$, for $n = k + \ell + 1$.

The $(u(1) \oplus \cdots \oplus u(1))^{(1)}$ classification has connections to rational points on Grassmannians. The Grassmannian is (essentially) the moduli space for the Narain compactifications of the (classical) lattice string. It would be very interesting to interpret other large families of physical invariants as special points on other moduli spaces.

These new connections relating various physical invariant classifications to other areas of math seem to indicate that although the physical invariant classifications are difficult, they could be well worth the effort and be of interest outside RCFT. Once the physical invariant lists are obtained, we will still have the fascinating task of explaining and developing all these mysterious connections. These thoughts keep me going!

Another motivation for completing these lists comes from their relation to subfactor theory in von Neumann algebras[18]. These algebras (see e.g. [22]) can be thought of as symmetries of a (generally infinite) group. Their building blocks are called *factors*. Jones initiated the combinatorial study of *subfactors* N of M (i.e. inclusions $N \subseteq M$ where M, N are factors), relating it to e.g. knots, and for this won a Fields medal in 1990. Jones assigned to each subfactor $N \subseteq M$ a numerical invariant called an 'index', a sort of (generally irrational) ratio of dimensions. Graphs (called principal and dual principal) are also associated to subfactors. A much more refined subfactor invariant, called a 'paragroup', has been introduced by Ocneanu. It is essentially equivalent to a (2+1)-dimensional topological field theory. Moreover, any RCFT can be assigned a paragroup, and any paragroup (via a process called asymptotic inclusion which is akin to Drinfeld's quantum doubling of Hopf algebras) yields an RCFT. See [22] for details.

Böckenhauer-Evans [4] have recently developed this much further, and have clarified the fusion graph ↔ physical invariant relation. The fusion graphs will correspond to subfactor principal graphs. In the work of Di Francesco-Petkova-Zuber, that relation seems to be only empirical (i.e. nonconceptual).

Subfactor theory together with singularity theory is our best hope at present for understanding and generalising the A-D-E meta-pattern.

---

[18] For reasons of necessity, in the following discussion I'll take more liberties than usual in the presentation.

Part 2. Monstrous Moonshine

2.1. INTRODUCTION

In 1978, John McKay made a very curious observation. One of the well-known[19] functions of classical number theory is the j-function[20], given by

$$\begin{aligned} j(\tau) :=& \frac{(1+240\sum_{n=1}^{\infty}\sigma_3(n)\,q^n)^3}{q\prod_{n=1}^{\infty}(1-q^n)^{24}} = \frac{\Theta_{E_8}(\tau)^3}{\eta(\tau)^{24}} \\ =& q^{-1}+744+196\,884\,q+21\,493\,760\,q^2+864\,299\,970\,q^3+\cdots \end{aligned} \tag{2.1.1}$$

Here as elsewhere in this paper, $q=\exp[2\pi\mathrm{i}\,\tau]$. Also, $\sigma_3(n)=\sum_{d|n}d^3$, Θ_{E_8} is the theta function of the E_8 root lattice, and η is the Dedekind eta. What is important here are the values of the first few coefficients. What McKay noticed was that $196\,884\approx 196\,883$. Closer inspection shows $21\,493\,760\approx 21\,296\,876$, and $864\,299\,970\approx 842\,609\,326$. In fact,

$$196\,884 = 196\,883+1 \tag{2.1.2a}$$

$$21\,493\,760 = 21\,296\,876+196\,883+1 \tag{2.1.2b}$$

$$864\,299\,970 = 842\,609\,326+21\,296\,876+2\cdot 196\,883+2\cdot 1 \tag{2.1.2c}$$

The numbers on the right-side are the dimensions of the smallest irreducible representations of the Monster finite simple group $\mathbb{M}$ (in 1978 it still wasn't certain that $\mathbb{M}$ even existed so back then these numbers were merely conjectural). The same game could be played with other coefficients of the j-function. With numbers so large, it seemed to him doubtful that this numerology was merely a coincidence. On the other hand, it was hard to imagine any deep conceptual connection between the Monster and the j-function: they seem completely unrelated.

In November 1978 he mailed the 'McKay equation' (2.1.2a) to John Thompson. At first Thompson dismissed this as nonsense, but after checking the next few coefficients he became convinced. He then added a vital piece to the puzzle. It should be well-known that when one sees a nonnegative integer, it often helps to try to interpret it as the dimension of some vector space. Essentially, that is what McKay was proposing here. (2.1.2) are really hinting that there is a 'graded' representation V of $\mathbb{M}$:

$$V = V_{-1}\oplus V_1\oplus V_2\oplus V_3\oplus\cdots$$

where $V_{-1}=\rho_0$, $V_1=\rho_1\oplus\rho_0$, $V_2=\rho_2\oplus\rho_1\oplus\rho_0$, $V_3=\rho_3\oplus\rho_2\oplus\rho_1\oplus\rho_1\oplus\rho_0\oplus\rho_0$, etc, where ρ_i are the irreducible representations of $\mathbb{M}$ (ordered by dimension), and that

$$j(\tau)-744 = \dim_q(V) := \dim(V_{-1})\,q^{-1}+\sum_{i=1}^{\infty}\dim(V_i)\,q^i\ , \tag{2.1.3}$$

---

[19] 'Well-known' is math euphemism for 'a basic result of which until recently we were utterly ignorant.' As Conway later said, "the j-function was 'well-known' to other people, but not 'well-known' to me."

[20] This and other technical terms used in this introduction will be carefully explained in the following subsections. This section is merely offered as a quick overview.

the graded dimension of V.

Thompson suggested that we twist $\dim_q(V)$, i.e. that more generally we consider the series (now called the *McKay-Thompson series*)

$$T_g(\tau) := \mathrm{ch}_{V,q}(g) = \mathrm{ch}_{V_{-1}}(g)\, q^{-1} + \sum_{i=1}^{\infty} \mathrm{ch}_{V_i}(g)\, q^i \ , \tag{2.1.4}$$

for each element $g \in \mathbb{M}$. The point is that, for any group representation ρ, the character value $\mathrm{ch}_\rho(id.)$ equals the dimension of ρ, and so $T_{id.}(\tau) = j(\tau) - 744$ and we recover (2.1.2) as special cases. But there are many other possible choices of $g \in \mathbb{M}$. Thompson couldn't guess what these functions T_g would be, but he suggested that they too might be interesting. This is a nice thought: when we see a positive integer, we should try to interpret it as a dimension of a vector space; if there is a symmetry present, then it may act on the vector space — i.e. our vector space may carry a representation of that symmetry group — in which case we can apply the Thompson trick and see what if any significance the other character values have in our context.

Conway and Norton [12] did precisely what Thompson asked. Conway called it "one of the most exciting moments in my life" [11] when he opened Jacobi's foundational (but 150 year old!) book on elliptic and modular functions and found that the first few terms of the McKay-Thompson series agreed perfectly with the first few terms of certain special functions, namely the Hauptmoduls of various genus 0 modular groups. Monstrous Moonshine was officially born.

The word 'moonshine' here is English slang for 'unsubstantial or unreal'. It was chosen by Conway to convey as well the feeling that things here are dimly lit, and that Conway-Norton were 'distilling information illegally' from the Monster character table.

In fact the first incarnation of Moonshine goes back to Andrew Ogg in 1975. He was in France describing his result that the primes p for which the group $\Gamma_0(p)+$ has genus 0, are $\{2, 3, 5, 11, 13, 17, 19, 23, 29, 31, 41, 47, 59, 71\}$. $\Gamma_0(p)+$ is the group generated by $\begin{pmatrix} 0 & 1 \\ -p & 0 \end{pmatrix}$ and $\Gamma_0(p)$, and is the normaliser of $\Gamma_0(p)$ in $\mathrm{SL}_2(\mathbb{R})$ (this sentence will make a little more sense after §2.3, but it isn't important here to understand it). He also attended a lecture by Jacques Tits, who was describing a newly conjectured simple group. When Tits wrote down the prime decomposition of the order of that group (see (2.2.1) below), Ogg noticed its prime factors precisely equalled his list of primes. Presumably as a joke, he offered a bottle of Jack Daniels' whisky to the first person to explain the coincidence.

The next step was accomplished by Griess in 1980, with the construction of the Monster[21] $\mathbb{M}$, and with it the proof that the conjectured character table for $\mathbb{M}$ was correct. Griess did this by explicitly constructing the 196883-dimensional representation ρ_1; it turns out to have a (commutative nonassociative) algebra structure, now called the *Griess algebra*. Though this paper was clearly important, the construction was artificial and 100 pages long: since the Monster is presumably a natural mathematical object (see §2.2), an elegant construction for it should exist. This was ultimately accomplished in the mid 1980s with the construction by Frenkel-Lepowsky-Meurman [23] of the Moonshine module

---

[21] Griess also came up with the symbol for the Monster; Conway came up with the name.

$V^\natural$ and its interpretation by Borcherds as a *vertex operator algebra*. The Griess algebra appears naturally in $V^\natural$, as we shall see. $V^\natural$ does indeed seem to be a 'natural' mathematical structure, and $\mathbb{M}$ is its automorphism group: in fact $V^\natural$ is the graded representation V of $\mathbb{M}$ conjectured by McKay and Thompson.

Connections with physics (CFT) go back to Dixon-Ginsparg-Harvey [19] in 1988, in a paper titled "Beauty and the beast: Superconformal symmetry in a Monster module". The Moonshine module $V^\natural$ can be interpreted as the string theory for a $\mathbb{Z}_2$-orbifold of free bosons compactified on the torus $\mathbb{R}^{24}/\Lambda_{24}$ (Λ_{24} is the Leech lattice). Many aspects of Moonshine make complete sense within CFT, but some (e.g. the genus zero property) remain more obscure. (Though in 1987 Moore speculated that the 0-genus of $\Gamma_0(a)+$ could be related to the vanishing of the cosmological constant in certain string theories related to $\mathbb{M}$, and Tuite [57] related genus-zero with the conjectured uniqueness of $V^\natural$.) Nevertheless this helps make the words of Dyson ring prophetic: "I have a sneaking hope, a hope unsupported by any facts or any evidence, that sometime in the twenty-first century physicists will stumble upon the Monster group, built in some unsuspected way into the structure of the universe" [21].

Finally, in 1992 Borcherds [5] completed the proof of the Conway-Norton conjectures by showing $V^\natural$ is the desired representation V. The full conceptual relationship between the Monster and the Hauptmoduls (like j) seems to remain 'dimly lit', although much progress has been realised. This is a subject where it is much easier to conjecture than to prove, and we are still awash in unresolved conjectures.

McKay also noticed in 1978 that similar coincidences hold if $\mathbb{M}$ and $j(\tau)$ are replaced with the Lie group $E_8(\mathbb{C})$ and $(qj(q))^{\frac{1}{3}} = 1 + 248q + \cdots$. This turns out to be much easier to explain, and in 1980 both Kac and Lepowsky remarked that the unique level 1 highest-weight representation of the affine algebra $E_8^{(1)}$ has graded dimension $(qj(q))^{\frac{1}{3}}$.

Moonshiners have a little chip on their shoulders. Modern math, they say, tends to be a little too infatuated with the pursuit of generalisations for generalisations' sake. Surely a noble goal for math is to find interesting and fundamentally new theorems. It can be argued that both history and common-sense suggest that to this end it is most profitable to look simultaneously at both exceptional structures and generic structures, to understand the special features of the former in the context of the latter, and to be led in this way to a new generation of exceptional and generic structures. Moonshiners would sympathise with those biologists who study the duck-billed platypus and lungfish rather than hide them in the closet as monsters: BECAUSE they appear to be unique, those animals presumably have much to teach us about our general understanding of evolution, etc.

It often seems to people that Moonshine can't be very deep: the Conway-Norton conjectures seem to be so finite[22] and specialised. There only are 171 distinct McKay-Thompson series T_g in Monstrous Moonshine, after all. The whole point though is to try to understand *why* the Monster and the Hauptmoduls are so related, and then to try to extend and apply this understanding to other contexts. Moonshine is still young, and our understanding remains incomplete. But already math has benefitted: e.g. we now have a natural definition of $\mathbb{M}$ (as the automorphism group of $V^\natural$), and Moonshine helped lead us to the rich structures of generalised Kac-Moody algebras and vertex operator algebras.

---

[22] Indeed the Moonshine conjectures *are* finite (it is enough to check the first 1200 coefficients), and a slightly weaker form was quickly proved on a computer by Atkin, Fong and Smith [54]. However this sort of argument adds no light to Moonshine, and tells us nothing of V except that it exists.

We will see that Moonshine involves the interplay between *exceptional structures* such as the number 24, the Leech lattice Λ_{24}, the Monster group $\mathbb{M}$, and the Moonshine module $V^\natural$, and *generic structures* such as modular functions, vertex operator algebras, generalised Kac-Moody algebras, and conformal field theories. The following sections will introduce the reader to many of these structures, as we use Moonshine as another happy excuse to take a second little tour through modern mathematics.

2.2. INGREDIENT #1: FINITE SIMPLE GROUPS AND THE MONSTER

A readable introduction to the basics of finite group representation theory is [25]. The finite simple groups are described in [33]; see also [11]. Group representations were introduced in §1.3.

A *normal* subgroup H of a group is one obeying $gHg^{-1} = H$ for all $g \in G$. These are important because the set G/H of 'cosets' gH has a natural group structure precisely when H is normal. Every group has two trivial normal subgroups: itself and $\{1\}$. If these are the only normal subgroups, the group is called *simple*. It is conventional to regard the trivial group $\{1\}$ as not simple (just as 1 is conventionally regarded as not prime). An alternate definition of a (finite) simple group G is that if $\varphi : G \to H$ is any group homomorphism (i.e. structure-preserving map: $\varphi(gg') = \varphi(g)\varphi(g')$), then φ is either constant (i.e. $\varphi(G) = \{1\}$), or φ is one-to-one.

The importance of simple groups is provided by the *Jordan-Hölder Theorem*. By a 'composition series' for a group G, we mean a nested sequence

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_k \supset H_{k+1} = \{1\}$$

of groups such that H_i is normal in H_{i-1}, and H_{i-1}/H_i (called a 'composition factor') is simple. Any finite group G has at least one composition series. If $H'_0 \supset \cdots \supset H'_{\ell+1} = \{1\}$ is a second composition series for G, then Jordan-Hölder says that $k = \ell$ and, up to a reordering π, the simple groups H_{i-1}/H_i and $H'_{\pi j-1}/H'_{\pi j}$ are isomorphic.

For example, the cyclic group $\mathbb{Z}_n$ of order(=size) n — you can think of it as the integers modulo n under addition — is simple iff n is prime. Consider the group $\mathbb{Z}_{12} = \langle 1 \rangle$. Two composition series are

$$\mathbb{Z}_{12} \supset \langle 2 \rangle \supset \langle 4 \rangle \supset \langle 0 \rangle$$
$$\mathbb{Z}_{12} \supset \langle 3 \rangle \supset \langle 6 \rangle \supset \langle 0 \rangle$$

corresponding to composition factors $\mathbb{Z}_2$, $\mathbb{Z}_2$, $\mathbb{Z}_3$, and $\mathbb{Z}_3$, $\mathbb{Z}_2$, $\mathbb{Z}_2$. Of course this is consistent with Jordan-Hölder. This is reminiscent of the fact that $2 \cdot 2 \cdot 3 = 3 \cdot 2 \cdot 2$ are both prime factorisations of 12.

There is some value to regarding finite groups as a massive generalisation of the notion of number. The number n can be identified with the cyclic group $\mathbb{Z}_n$. The divisor of a number corresponds to a normal subgroup, so a prime number corresponds to a simple group. The Jordan-Hölder Theorem generalises the uniqueness of prime factorisations. That you can build up any number by multiplying primes, is generalised to building up a group by semi-direct products (more generally, by group extensions): if H is a normal subgroup of G, then G will be an extension of H by the quotient group G/H.

Note however that $\mathbb{Z}_6 \times \mathbb{Z}_2$ and $\mathfrak{S}_3 \times \mathbb{Z}_2$ — both different from $\mathbb{Z}_{12}$ — will also have $\mathbb{Z}_2, \mathbb{Z}_2, \mathbb{Z}_3$ as composition factors: unlike for numbers, 'multiplication' here does not give a unique answer. The semidirect product $\mathbb{Z}_2 \rtimes \mathbb{Z}_2$ can equal either $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$, depending on how the product is taken. More precisely, the notation $G \rtimes G'$ means a group where every element can be written uniquely as a pair (g, g'), for $g \in G$ and $g' \in G'$, and where the group operation is $(g, g')(h, h') = (\text{stuff}, gh)$.

Thus simple groups have an importance for group theory approximating what primes have for number theory. One of the greatest accomplishments of twentieth century math is surely the classification of the finite simple groups. (On the other hand, group extensions turn out to be technically quite difficult and leads one into group cohomology.) This work, completed in the early 1980s (although gaps are continually being discovered and filled in the arguments), runs to approximately 15 000 journal pages, spread over 500 individual papers, and is the work of a whole generation of group theorists. A modern revision is currently underway (see e.g. [34]) to simplify the proof and find and fill all gaps, but the final proof is still expected to be around 4000 pages long. The resulting list is:

- the cyclic groups $\mathbb{Z}_p$ (p a prime);
- the alternating groups $\mathfrak{A}_n$ for $n \geq 5$;
- 16 families of Lie type;
- 26 sporadics.

We've already met the cyclic groups. The alternating group $\mathfrak{A}_n$ consists of the even permutations in the symmetric group $\mathfrak{S}_n$, and so has order(=size) $\frac{1}{2}\,n!$. The groups of Lie type are essentially Lie groups defined over finite fields[23] $\mathbb{F}_q$ (such as $\mathbb{Z}_p$), sometimes 'twisted' in certain senses. The simplest example is $\mathrm{PSL}_n(\mathbb{F}_q)$, which consists of the $n \times n$ matrices with entries in $\mathbb{F}_q$, with determinant 1, quotiented out by the centre of $\mathrm{SL}_n(\mathbb{F}_q)$ (namely the scalar matrices $\mathrm{diag}(a, a, \ldots, a)$ for $a^n = 1$) (except for $\mathrm{PSL}_2(\mathbb{Z}_2)$ and $\mathrm{PSL}_2(\mathbb{Z}_3)$, which aren't simple).

Note that the determinant $|\rho(g)|$ for any representation ρ of any (noncyclic) simple group must be 1, otherwise we would violate the homomorphism definition of simple group (try to see why). Also, the centre of any (noncyclic) simple group must be trivial (why?). The smallest noncyclic simple group is $\mathfrak{A}_5$, with order 60.[24] It is the same as (isomorphic to) $\mathrm{PSL}_2(\mathbb{Z}_5)$ and $\mathrm{PSL}_2(\mathbb{F}_4)$, and can also be expressed as the group of all rotations (reflections have determinant -1 and so cannot belong to any simple group) of $\mathbb{R}^3$ that bring a regular icosahedron back to itself.

---

[23] There is a finite field with q elements, iff q is a power of a prime. For each such q, there is only 1 field of that size. The field with prime p elements is the integers taken mod p.

[24] This implies, incidentally, that if G and H are any two groups with the same order below 60, then they will have the same composition factors.

The smallest sporadic group is the Mathieu group M_{11}, order 7920, discovered in 1861[25]. The largest is the Monster $\mathbb{M}$, conjectured by Fischer and Griess in 1973 and finally proved to exist by Griess in 1980. Its order is[26]

$$\|\mathbb{M}\| = 2^{46} \cdot 3^{20} \cdot 5^{9} \cdot 7^{6} \cdot 11^{2} \cdot 13^{3} \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \times 10^{53} \ . \quad (2.2.1)$$

20 of the 26 sporadics are involved in (i.e. are quotients of subgroups of) the Monster. Some relations among $\mathbb{M}$, the Leech lattice Λ_{24} and the largest Mathieu group M_{24} are given in Chapters 10 and 29 of [13].

Moonshine hints at a tantalising connection between the classification of finite simple groups, and the classification of RCFTs discussed in Part 1. Speculates [23] (page xli): "One can certainly hope for a uniform description of the finite simple groups as automorphism groups of certain vertex operator algebras — or conformal quantum field theories. If such a quantum field theory could somehow be attached a priori to a finite simple group, the classification of such theories, a problem of great current interest among string theorists, might some day be part of a new approach to the classification of the finite simple groups. On the other hand, can the known classification of the finite simple groups help in the classification of conformal field theories?"

2.3. INGREDIENT #2: MODULAR FUNCTIONS AND HAUPTMODULS

A readable introduction to some of the topics discussed in this section is [16,42,60].

We know from complex analysis that the group $\mathrm{SL}_2(\mathbb{R})$ of 2×2 matrices with real entries and determinant 1, acts on the upper-half plane $\mathcal{H} = \{\tau \in \mathbb{C} \,|\, \mathrm{Im}(\tau) > 0\}$ by fractional linear (or *Möbius*) transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau := \frac{a\tau + b}{c\tau + d} \ . \quad (2.3.1)$$

For example the matrix $S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponds to the function $\tau \mapsto -1/\tau$, while the matrix $T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ corresponds to the translation $\tau \mapsto \tau + 1$. Since $\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ correspond to the same Möbius transformation, strictly speaking our group here is $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm I\}$.

The only reason this action (2.3.1) of the 2×2 matrices on complex numbers (or more precisely the Riemann sphere $\mathbb{C} \cup \{\infty\}$) might not look strange to us, is because familiarity breeds numbness. What we really have is a natural action of $n \times n$ matrices on $\mathbb{C}^n$, and this induces their action on $\mathbb{C}^{n-1}$ (together with a codimension-2 set of 'points at infinity') by interpreting $\mathbb{C}^n$ as projective coordinates for $\mathbb{C}^{n-1}$. Specialising to $n = 2$ gives us (2.3.1). In projective geometry, 'parallel lines' intersect at ∞. Projective coordinates allow one to treat 'finite' and 'infinite' points on an equal footing.

---

[25] Although his arguments apparently weren't very convincing. In fact some people, including the Jordan of Jordan-Hölder fame, argued in later papers that the largest of Mathieu's sporadic groups couldn't exist.

[26] The inquisitive reader, hungry for more 'coincidences', may have noticed that 196883 and 21296876 — see (2.1.2) — exactly divide the order of the Monster. Indeed this will hold for any finite group: the dimensions of the irreducible representations of a finite group will always divide its order.

Consider $\Gamma := \mathrm{SL}_2(\mathbb{Z})$, the subgroup of $\mathrm{SL}_2(\mathbb{R})$ consisting of the matrices with integer entries. It can be shown that it is generated by S and T (in other words, every matrix $\alpha \in \Gamma$ can be expressed as a monomial in S and T). For reasons that will be clear shortly, consider the extended upper-half plane $\overline{\mathcal{H}} := \mathcal{H} \cup \{\mathrm{i}\infty\} \cup \mathbb{Q}$ — the extra points $\{\mathrm{i}\infty\} \cup \mathbb{Q}$ are called *cusps*. Γ acts on $\overline{\mathcal{H}}$ (e.g. S interchanges 0 and $\mathrm{i}\infty$). By a *modular function* for Γ, we mean a meromorphic function $f : \overline{\mathcal{H}} \to \mathbb{C}$, symmetric with respect to Γ: i.e. $f(\alpha(\tau)) = f(\tau)$ for all $\alpha \in \Gamma$. Note that we require f to be meromorphic at the cusps (e.g. polynomials are meromorphic at $\mathrm{i}\infty$, but e^z is not).

It is not obvious why modular functions should be interesting, but in fact they are one of the most fundamental notions in modern number theory (see the last paragraph of §1.6). For example, consider the question of writing numbers as sums of squares. We can write $5 = 1^2 + (-2)^2 = (-1)^2 + 1^2 + 0^2 + 1^2 + (-1)^2$, to give a couple of trivial examples. Let $N_n(k)$ be the number of ways we can write the integer n as a sum of k squares, counting order and signs. For example $N_5(1) = 0$ (since 5 is not a perfect square), $N_5(2) = 8$ (since $5 = (\pm 1)^2 + (\pm 2)^2 = (\pm 2)^2 + (\pm 1)^2$), $N_5(3) = 24$, etc. Their generating functions are:[27]

$$\sum_{n=0}^{\infty} N_n(k)\, q^n = (\theta_3(q))^k \ ,$$

where

$$\theta_3(q) = 1 + 2q + 2q^4 + \cdots = \sum_{n \in \mathbb{Z}} q^{n^2}$$

is called a *theta function*. It turns out that θ_3 transforms nicely with respect to Γ, once we make the change-of-variables $q = \exp[\pi \mathrm{i} \tau]$. This takes work to show. For example, θ_3 is clearly invariant under the action of $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, and a little work (from e.g. Poisson summation) shows that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ takes $\theta_3(\tau)$ to $\sqrt{\frac{\tau}{\mathrm{i}}}\,\theta_3(\tau)$. θ_3 is not precisely a modular function (it is a 'modular form of weight $\frac{1}{2}$' for $\Gamma_0(4)$), but this simple example illustrates the point that Γ (and related groups) appear throughout number theory. More on this shortly.

That important change-of-variables $q = \exp[\pi \mathrm{i} \tau]$ was introduced by Jacobi early last century, in his analysis of 'elliptic integrals'. The theory is beautiful and poorly remembered today, which is very disappointing considering how much of modern math was touched by it. I strongly recommend the book [10], written over a century ago; the style and motivation of math in our century is different from that in Jacobi's, and we've lost a little in motivation what we've gained in power. I'll briefly sketch Jacobi's theory.

---

[27] A fundamental principle in math is: whenever you have a subscript with an infinite range, make a power series (called a *generating function*) out of it.

Just as we could develop a theory of 'circular functions' (i.e. sine etc.) starting from the integral $s(a) = \int_0^a \frac{dx}{\sqrt{1-x^2}}$, so can we develop a theory of 'elliptic functions' starting from the 'elliptic integral' $F(k,a) = \int_0^a \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$. Inverting $s(a)$ gives a function both more useful and with nicer properties than $s(a)$: we call it $\sin(u)$. Similarly, for any k the elliptic function $\mathrm{sn}(k,u)$ is defined by $u = F(k, \mathrm{sn}(k,u))$. Just as we can define a numerical constant π by $\sin(\frac{1}{2}\pi) = 1$ (i.e. $\frac{1}{2}\pi = \int_0^1 \frac{dx}{\sqrt{1-x^2}}$), we get a function $K(k) = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$. Just as $\sin(u)$ has period $4(\frac{1}{2}\pi)$, so has sn u-period $4K(k)$. sn also turns out to have u-period $4\mathrm{i}\,K(k')$ where $k' = \sqrt{1-k^2}$ — today we take this as the starting point and define an elliptic function to be doubly periodic (see [42] or Cohen in [60]).

The theta functions aren't elliptic functions, but they are closely related, and e.g. sn can be written as a quotient of them. In Jacobi's language, we have

$$\theta_3(\frac{\mathrm{i}\,K(k')}{K(k)}) = \sqrt{\frac{2K(k)}{\pi}}\ .$$

The 'modular transformation' $\tau \mapsto \frac{-1}{\tau}$ corresponds to interchanging the 'modulus' k with the 'complementary modulus' k', and thus is completely natural in Jacobi's theory. The important formula $\theta_3(\frac{-1}{\tau}) = \sqrt{\frac{\tau}{\mathrm{i}}}\,\theta_3(\tau)$ is trivial here.

A certain interpretation of modular functions also indicates their usefulness, and played an important role in Part 1. A *torus* is something that looks like the surface of a bagel, at least as far as its topology is concerned. For example, the Cartesian product $S^1 \times S^1$ of circles is a torus (think of one circle being the contact-circle of the bagel with the table on which it rests, then from each point on that horizontal circle imagine placing a vertical circle perpendicular to it, like a rib; together all these ribs fill out the bagel's surface). A more sophisticated example of a torus is an elliptic curve (a complex curve of the form $y^2 = ax^3 + bx^2 + cx + d$ and a special point on it playing the role of 0). A final example is the quotient $\mathbb{C}/\Lambda$ of the complex plane $\mathbb{C}$ with a 2-dimensional lattice Λ (we saw lattices in §1.6; Λ here will be a discrete doubly-periodic set of points in $\mathbb{C}$, containing 0). It turns out that certain equivalence classes of tori (e.g. with respect to conformal or complex-analytic equivalence) always contain a representative torus of the form $\mathbb{C}/\Lambda$, where Λ consists of all points $\mathbb{Z} + \mathbb{Z}\,\tau$, for some $\tau \in \mathcal{H}$. (Incidentally, the cusps correspond to degenerate tori.) In other words, these equivalence classes are parametrized by complex numbers τ in $\mathcal{H}$. So if we have a complex-valued function F on the set of all tori, which is e.g. conformally invariant (an example is the genus-one partition function $\mathcal{Z}$ in conformal field theories — see §1.1), then we can consider F as a well-defined function $F : \mathcal{H} \to \mathbb{C}$. However, it turns out that different points τ in $\mathcal{H}$ correspond to the same equivalence class of tori: e.g. the lattice for τ is the same as that for $\tau + 1$, and these are a rescaling of that for $-1/\tau$. Thus $F(\tau) = F(\tau+1) = F(-1/\tau)$, because $\tau, \tau+1, -1/\tau$ all represent equivalent tori. Since $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$ generate $\mathrm{PSL}_2(\mathbb{Z})$, what in fact we find is that F has Γ as its group of symmetries. One often says that Γ is the 'modular group of the torus', and that the orbit space $\Gamma\backslash\mathcal{H}$ is the 'moduli space' of (conformal equivalence classes of) tori. $\mathcal{H}$ is called its 'Teichmüller space' or 'universal cover'. This is exactly analogous to $S^1 = \mathbb{R}/\mathbb{Z}$: $\mathbb{R}$ is its universal cover and $\mathbb{Z}$ is its 'modular group' (or 'mapping class group'). Another example: the Teichmüller space for (conformal equivalence classes of) 'pair-of-pants', or equivalently a disc minus two open interior disks, is $\mathbb{R}^3$ (an ordered triple), while its modular group is the symmetric group $\mathfrak{S}_3$ and its moduli space consists of unordered triples. Incidentally, we write $\Gamma\backslash\mathcal{H}$ instead of $\mathcal{H}/\Gamma$ because the group Γ acts on $\mathcal{H}$ 'on the left'. A good introduction to the geometry here is [55].

In any case, a surprising number of innocent-looking questions in number theory can be dragged (usually with effort) into the richly developed realm of elliptic curves and modular functions, and it is there they are often solved. For instance, we all know the ancient Greeks were interested in Pythagorean triples: find all integer solutions a, b, c to $a^2 + b^2 = c^2$, i.e. find all integer (or if you prefer, rational) right-angle triangles. They solved this by elementary means: choose any integers (or rationals) x, y and put $u = \frac{x^2-y^2}{x^2+y^2}$, $v = \frac{2xy}{x^2+y^2}$; then $u^2 + v^2 = 1$ and (multiplying by the denominator) this gives all Pythagorean triples.

There are two ways of extending this problem. One is to ask which $n \in \mathbb{Z}$ can arise as areas of these rational right-angle triangles. It turns out $n = 5$ is the smallest one: $a = \frac{3}{2}$, $b = \frac{20}{3}$, $c = \frac{41}{6}$ works ($5 = \frac{1}{2}(\frac{3}{2})(\frac{20}{3})$ and $(\frac{3}{2})^2 + (\frac{20}{3})^2 = (\frac{41}{6})^2$). This is a hard problem — just try to show $n = 1$ cannot work. $n = 157$ turns out to work: the simplest triangle has a and b as quotients of integers of size around 10^{25}, and c as the quotient of integers around 10^{47}. Although this problem was studied by the ancient Greeks and also by the Arabs in the 10th century, it was finally cracked in the 1980s. It was solved by first translating it into the question of whether the elliptic curve $y^2 = x^3 - n^2x$ has infinitely-many rational points, and then applying all the rich 20th century machinery to answering that question.

The other continuation of the Pythagorean triples question is more famous: find all integer solutions to $a^n + b^n = c^n$ (or equivalently all rational solutions to $a^n + b^n = 1$). 350 years ago Fermat wrote in the margin of the book he was reading (the book was describing the Greek solution to Pythagorean triples) that he had found a "truly marvelous" proof that for $n > 2$ there are no nontrivial solutions, but that the margin was too narrow to contain it. This result came to be known as 'Fermat's Last Theorem'[28] and despite considerable effort no one has succeeded in rediscovering his proof. Most people today believe that Fermat soon realised his 'proof' wasn't valid, otherwise he would have alluded to it in later letters. In any case, a very long and complicated proof was finally achieved in the 1990s: the 'Taniyama conjecture' says that a certain function associated to any elliptic curve over $\mathbb{Q}$ will be modular; if $a^n + b^n = c^n$ for some $n > 2$, then the elliptic curve $y^2 = x^3 + (a^n - b^n)x^2 - a^nb^n$ will violate the Taniyama conjecture; finally Wiles proved the Taniyama conjecture is true.

To most mathematicians, the 'area-n problem' and 'Fermat's Last Theorem' are interesting only because they can be related to elliptic curves and modular forms — it's easy to ask hard questions in math, but most questions tend to be stale. Number theory is infatuated with modular stuff because (in increasing order of significance) (a) it's exceedingly rich, with lots of connections to other areas of math and math phys; (b) it's a battleground on which many innocent-looking but hard-to-crack problems can be slain; and (c) last generation's number theorists also worked on modular stuff.

---

[28] It was called his 'Last Theorem' because it was the last of his 48 margin notes to be proved by other mathematicians — another one is discussed in Section 1.8. The story of Fermat's Last Theorem is a fascinating one, but alas this footnote is too small to do it credit. See for instance the excellent book [53].

In any case, modular functions turn out to be important for math (and mathematical physics) even though they may at first glance look artificial. Poincaré explained how to study them. He said to look at the orbits of $\overline{\mathcal{H}}$ with respect to Γ. For example, one orbit, hence one point in $\Gamma\backslash\overline{\mathcal{H}}$, contains all cusps. We write this as $\Gamma\backslash\overline{\mathcal{H}}$, and give it the natural topological structure (i.e. 2 points $[\tau], [\tau'] \in \Gamma\backslash\overline{\mathcal{H}}$ are considered 'close' if the 2 sets $\Gamma\tau, \Gamma\tau'$ nearly overlap). Note first that by applying T repeatedly, every point in $\mathcal{H}$ corresponds to a point in the vertical strip $-\frac{1}{2} \leq \mathrm{Re}(\tau) \leq \frac{1}{2}$ — in fact to a unique point in that strip, if we avoid the two edges. S is an inversion through the unit circle, so it permits us to restrict to those points in the vertical strip which are distance at least 1 from the origin. The resulting region R is called a fundamental region for Γ. Apart from the boundary of R, every Γ-orbit will intersect R in one and only one point.

What should we do about the boundary? Well, the edge $\mathrm{Re}(\tau) = -\frac{1}{2}$ gets mapped by T to the edge $\mathrm{Re}(\tau) = \frac{1}{2}$, so we should identify (=glue together) these. The result is a cylinder running off to infinity, with a strange lip at the bottom. S tells us how we should close that lip: identify $\mathrm{i}e^{\mathrm{i}\theta}$ and $\mathrm{i}e^{-\mathrm{i}\theta}$. This seals the bottom of the cylinder, so we get an infinitely tall cup with a strangely puckered base. In fact the top of this cup is also capped off, by the cusp $\mathrm{i}\infty$. So what we have (topologically speaking) is a *sphere.* It does not look like a smooth sphere, but in fact it inherits the smoothness of $\mathcal{H}$.

Incidentally, topological manifolds of dimension ≤ 3 always have a unique compatible smooth structure. 'Topological structure' means you can speak of continuity or closeness, 'smooth structure' means you can also do calculus. On the other hand, $\mathbb{R}^4$ has infinitely many smooth structures compatible with its topological structure; mysteriously, all other Euclidean spaces $\mathbb{R}^n$ have a unique smooth structure! Thus both mathematics and physics single out 4-space. Coincidence???

So anyways, what this construction of $\Gamma\backslash\overline{\mathcal{H}}$ means is that a modular function can be reinterpreted as a meromorphic complex-valued function on this sphere. This is very useful, because our undergraduate complex variables class taught us all about meromorphic complex-valued functions f on the Riemann sphere $\mathbb{C} \cup \infty$. There are many meromorphic functions on $\mathbb{C}$, but to also be meromorphic at ∞ forces f to be *rational*, i.e. $f(w) = \frac{\text{some polynomial } P(w)}{\text{some polynomial } Q(w)}$, where w is the complex parameter on the Riemann sphere. So our modular function $f(\tau)$ will simply be some rational function P/Q evaluated at the change-of-variables function $w = c(\tau)$ which maps us from our sphere $\Gamma\backslash\overline{\mathcal{H}}$ to the Riemann sphere. There are many different choices for this function $c(\tau)$, but the standard one is $c(\tau) = j(\tau)$, the j-function of (2.1.1)[29]. Thus, any modular function can be written as a rational function $f(\tau) = P(j(\tau))/Q(j(\tau))$ in the j-function. Conversely, any such function will be modular.

This is analogous to saying that any function $g(x)$ periodic under $x \mapsto x + 1$ can be thought of as a function on the unit circle $S^1 \subset \mathbb{C}$ evaluated at the change-of-variables function $x \mapsto e^{2\pi \mathrm{i}x}$, and hence has a Fourier expansion $\sum_n g_n \exp[2\pi \mathrm{i}nx]$.

---

[29] Historically, j was the standard choice, but in Moonshine the preferred choice would be the function $J = j - 744$ with zero constant term.

We can generalise this argument. Consider a subgroup G of $\mathrm{SL}_2(\mathbb{R})$ which is both not too big, and not too small. 'Not too big' means it should be *discrete*, i.e. the matrices in G can only get so close to the identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. To make sure G is 'not too small', it is enough to require that G contains some subgroup of the form

$$\Gamma_0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,|\, c \equiv 0 \pmod{N} \} \ , \tag{2.3.2a}$$

i.e. G must contain all matrices in Γ whose bottom-left entry is a multiple of N. So G must contain T, for example. We will also be interested only in those G which obey

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in G \ \Rightarrow \ t \in \mathbb{Z} \tag{2.3.2b}$$

i.e. the only translations in G are by integers. We will call a function $f : \overline{\mathcal{H}} \to \mathbb{C}$ a *modular function for* G if it is meromorphic (including at the cusps $\mathbb{Q} \cup \{\mathrm{i}\infty\}$), and if also f is symmetric with respect to G: $f \circ \alpha = f$ for all $\alpha \in G$. This implies we will be able to expand f as a Laurent series in q. We analyse this as before: look at the orbit space $\Sigma = G\backslash\overline{\mathcal{H}}$; because G is not too big, Σ will be a (Riemann) surface; because G is not too small, Σ will be compact.

The compact Riemann surfaces have been classified (up to homeomorphism — i.e. considering only topology as relevant), and are characterised by a number called the *genus*. Genus 0 is a sphere, genus 1 is a torus, genus 2 is like two tori resting side-by-side, etc. For example, the surface of a wine glass, or a fork, is topologically a sphere, while a coffee cup and a key will (usually) be tori. Eye glasses with the lenses popped out is a 2-torus, while a ladder with n rungs on it has genus $n-1$.

We will call G 'genus g' if its surface Σ has genus g. For example, $G = \Gamma_0(2)$ and $G = \Gamma_0(25)$ are both genus 0, while $\Gamma_0(50)$ is genus 2 and $\Gamma_0(24)$ is genus 3. Once again, we are interested here in the genus 0 case. As before, this means that there is a change-of-variables function we'll denote J_G which has the property that it's a modular function for G, and all other modular functions for G can be written as a rational function in it. Because of (2.3.2), we can choose J_G to look like

$$J_G(\tau) = q^{-1} + a_1(G)\, q + a_2(G)\, q^2 + \cdots$$

So J_G plays exactly the same role for G that $J := j - 744$ plays for Γ. J_G is called the Hauptmodul for G. (Incidentally for genus > 0, two generators, not one, are needed.)

For example, $\Gamma_0(2)$, $\Gamma_0(13)$ and $\Gamma_0(25)$ are all genus 0, with Hauptmoduls

$$J_2(\tau) = q^{-1} + 276\, q - 2048\, q^2 + 11202\, q^3 - 49152\, q^4 + 184024\, q^5 + \cdots \tag{2.3.3}$$

$$J_{13}(\tau) = q^{-1} - q + 2\, q^2 + q^3 + 2\, q^4 - 2\, q^5 - 2\, q^7 - 2\, q^8 + q^9 + \cdots \tag{2.3.4}$$

$$J_{25}(\tau) = q^{-1} - q + q^4 + q^6 - q^{11} - q^{14} + q^{21} + \cdots \tag{2.3.5}$$

The smaller the modular group, the smaller the coefficients of the Hauptmodul. In this sense, the j-function is optimally bad among the Hauptmoduls: e.g. for it $a_{23} \approx 10^{25}$.

An obvious question is, how many genus 0 groups (equivalently, how many Hauptmoduls) are there? It turns out that $\Gamma_0(p)$ is genus 0, for a prime p, iff $p-1$ divides 24. Thompson in 1980 proved that for any g, there are only finitely many genus g groups obeying our two conditions (2.3.2). In particular this means there are only finitely many Hauptmoduls. Over 600 Hauptmoduls with integer coefficients $a_i(G)$ are presently known.

2.4. THE MONSTROUS MOONSHINE CONJECTURES

We are now ready to make precise the main conjecture of Conway and Norton [12]. (We should emphasise though that there have been several other conjectures, some of which turned out to be partially wrong.)

They conjectured that for each element g of the Monster $\mathbb{M}$, there is a Hauptmodul

$$J_g(\tau) = q^{-1} + \sum_{n=1}^{\infty} a_n(g)\, q^n \tag{2.4.1}$$

for a genus 0 group G_g such that each coefficient $a_n(g)$ is an integer, and for each n the map $g \mapsto a_n(g)$ is a character of $\mathbb{M}$. They also conjectured that G_g contains $\Gamma_0(N)$ as a normal subgroup, for some N depending on the order of g.

Another way of saying this is that there exists an infinite-dimensional graded representation $V = V_{-1} \oplus \bigoplus_{n=1}^{\infty} V_n$ of $\mathbb{M}$ such that the McKay-Thompson series $T_g(\tau)$ in (2.1.4) is a Hauptmodul.

There are around 8×10^{53} elements to the Monster, so naively we may expect around 8×10^{53} different Hauptmoduls $J_g = T_g$. However the character of a representation evaluated at g and at hgh^{-1} will always be the same, so $J_g = J_{hgh^{-1}}$. Hence the relevant quantity is the number of conjugacy classes, which for $\mathbb{M}$ is only 194. Moreover, a character evaluated at g^{-1} will always be the complex conjugate of its value at g, but here all character values $\chi_{V_n}(g)$ are integers (according to the conjecture). Thus $J_g = J_{g^{-1}}$. The total number of distinct Hauptmoduls J_g arising in Monstrous Moonshine turns out to be only 171.

For example, if we choose g to be the identity, we recover $T_{id.} = J$. It turns out that there are precisely 2 different conjugacy classes of order 2 elements, one of them giving the Hauptmodul J_2 in (2.3.3). Similarly for 13, but J_{25} doesn't correspond to any conjugacy class of $\mathbb{M}$.

Moonshine provides an explanation for a forgotten mystery of classical mathematics: why are the coefficients of the j-function *positive* integers? On the other hand, that they are *integers* has long been important to number theory (complex multiplication, class field theory — see e.g. [16]).

There are lots of other less important conjectures. One which played a role in ultimately proving the main conjecture involves the *replication formulae*. Conway-Norton want to think of the Hauptmoduls J_g as being intimately connected with $\mathbb{M}$; if so, then the group structure of $\mathbb{M}$ should somehow directly relate different J_g. In particular, consider the power map $g \mapsto g^p$. Now, it was well-known that $j(\tau)$ has the property that $j(p\tau) + j(\frac{\tau}{p}) + j(\frac{\tau+1}{p}) + \cdots + j(\frac{\tau+p-1}{p})$ equals a polynomial in j, for any prime p (*sketch* *of proof:* it's a modular function for Γ, and hence equals a rational function of j; since its only poles will be at the cusps, the denominator polynomial must be trivial). Hence the same will hold for J. Explicitly we get

$$J(2\tau) + J(\frac{\tau}{2}) + J(\frac{\tau+1}{2}) = J^2(\tau) - 2a_1 \tag{2.4.2a}$$

$$J(3\tau) + J(\frac{\tau}{3}) + J(\frac{\tau+1}{3}) + J(\frac{\tau+2}{3}) = J^3(\tau) - 3a_1 J(\tau) - 3a_2 \tag{2.4.2b}$$

where $J(\tau) = \sum_k a_k q^k$. Slightly more complicated formulas hold in fact for any composite n. Conway and Norton conjectured that these formulas have an analogue for the Moonshine functions J_g in (2.4.1). In particular, (2.4.2) become for any $g \in \mathbb{M}$

$$J_{g^2}(2\tau) + J_g(\frac{\tau}{2}) + J_g(\frac{\tau+1}{2}) = J_g^2(\tau) - 2a_1(g) \tag{2.4.3a}$$

$$J_{g^3}(3\tau) + J_g(\frac{\tau}{3}) + J_g(\frac{\tau+1}{3}) + J_g(\frac{\tau+2}{3}) = J_g^3(\tau) - 3a_1(g)\, J_g(\tau) - 3a_2(g). \tag{2.4.3b}$$

These are examples of the replication formulae.

'Replication' concerns the power map $g \mapsto g^n$ in $\mathbb{M}$. Can Moonshine see more of the group structure of $\mathbb{M}$? A step in this direction was made by Norton [47], who associated a Hauptmodul to commuting elements g, h in $\mathbb{M}$. Physically [19], this corresponds to orbifold traces, i.e. the $V^\natural$ RCFT with boundary conditions twisted by g and h in the 'time' and 'space' directions. Still, we would like to see more of $\mathbb{M}$ in Moonshine.

An important part of the Monstrous Moonshine conjectures came a few years after [12]. Frenkel-Lepowsky-Meurman [23] constructed a graded infinite-dimensional representation $V^\natural$ of $\mathbb{M}$ and conjectured (correctly) that it is the representation in (2.1.4). $V^\natural$ has a very rich algebraic structure, which will be discussed in §2.6.

A major claim of [23] was that $V^\natural$ is a 'natural' structure (hence their notation). To see what they mean by that, it's best to view another simpler example of a natural construction: that of the *Leech lattice* Λ_{24}. Recall the discussion of (root) lattices in §1.6.

Λ_{24} is one of the most interesting lattices, and is related to Moonshine. It can be defined using 'laminated lattices'. Start with the 0-dimensional lattice $\Lambda_0 = \{0\}$, which consists of just a single point. Use it to construct a 1-dimensional lattice, with minimal (nonzero) norm 4, built out of infinitely many copies of Λ_0 laid side by side. The result of course is simply the even integers $2\mathbb{Z}$, which we will call here Λ_1. Now construct a 2-dimensional lattice, of minimum norm 4, built out of infinitely many copies of Λ_1 laid next to each other. There are lots of ways to do this, but choose the densest lattice possible. The result is unique: it is the hexagonal lattice A_2 scaled by a factor of $\sqrt{2}$: call it Λ_2. Continue in this way: Λ_3, Λ_4, Λ_5, Λ_6, Λ_7, and Λ_8 will be the root lattices A_3, D_4, D_5, E_6, E_7 and E_8, respectively, all scaled by $\sqrt{2}$. See [13] chapter 6 for a more complete treatment of laminated lattices.

The 24th repetition of this construction yields the Leech lattice. It is the unique 24-dimensional self-dual lattice with no norm-2 vectors, and provides among other things the densest known packing of 23-dimensional spheres in $\mathbb{R}^{24}$. Many of its properties are discussed throughout [13]. So lamination provides us with a sort of no-input construction of the Leech lattice, and a good example of the mathematical meaning of 'natural'. After dimension 24, it seems chaos results from the lamination procedure (there are 23 different 25-dimensional lattices that have an equal right to be called Λ_{25}, and over 75 000 are expected for Λ_{26}).

It is natural to ask about Moonshine for other groups. There is a partial Moonshine for the Mathieu groups M_{24} and M_{12} (which have about 2×10^8 and 10^5 elements resp.), the automorphism group .0 of Λ_{24} (which has about 8×10^{18} elements), and a few others — see e.g. [49]. These groups are either simple or almost simple (e.g. .0 is the direct product of $\mathbb{Z}_2$ with the simple group .1). More generally, there will be some sort of Moonshine for any group which is the automorphism group of a vertex operator algebra; the finite simple groups of Lie type should be automorphism groups of VOAs closely related to the affine algebras except defined over fields like $\mathbb{Z}_p$.

There is a geometric side to Moonshine, associated to names like Lian-Yau and Hirzebruch. In particular, Hirzebruch's 'prize question' asks for the construction of a 24-dimensional manifold on which $\mathbb{M}$ acts, whose twisted elliptic genus are the McKay-Thompson series. This is still open.

It should be emphasised that Monstrous Moonshine is a completely unexpected connection between finite groups and modular functions. Although there has been enormous progress in our understanding of this connection (so much so, that Richard Borcherds won the 1998 Fields medal for his work on this), there still is mystery at its heart. In particular, that $\mathbb{M}$ is associated with *modular functions* can be explained mathematically by it being the automorphism group of the Moonshine VOA $V^\natural$, and physically by the associated RCFT, but what is so special about $\mathbb{M}$ that these modular functions should be genus 0? We will come back to this in §2.9.

2.5. Formal Power Series

Vertex (operator) algebras (VOAs) are a mathematically precise formulation of the notion of *W-algebra* or *chiral algebra*[30] which is so central to conformal field theory (see §1.1). VOAs were first defined by Borcherds, and their theory has since been developed by a number of people (Frenkel, Lepowsky, Meurman, Zhu, Dong, Li, Mason, Huang, ...). Because our primary motivation here is Moonshine, I will only focus on one aspect of their theory (the connection with Lie algebras). Useful to consult while reading this review are the notes [27] — they take a more analytic approach to many of the things we discuss, and their approach (namely that of CFT) motivates beautifully much of VOA theory.

In quantum field theory the basic object is the quantum field, which roughly speaking is a choice of operator $\hat{A}(x)$ at each space-time point x. 'Operator' means something that 'operates on' functions or vectors. E.g. an indefinite integral is an operator, as is a derivative. The operators in the QFT act on the space spanned by the states $|\star\rangle$, and together form an infinite-dimensional vector space (e.g. a C^* algebra) — this infinite-dimensionality of QFT is a major source of its mathematical difficulties, and QFT still has not been put on completely satisfactory mathematical grounds.

---

[30] An alternate (and much more complicated) mathematical formulation of chiral algebra is due to Beilinson and Drinfeld, and belongs to algebraic geometry. See [28] for a good — but still difficult — review.

But another difficulty is that the quantum field $\hat{A}$ really isn't an operator-valued *function* of space-time. 'Function' is too narrow a concept. For example, one of the most familiar 'functions' in quantum mechanics is the Dirac delta $\delta(x)$. You see it for example in the canonical commutation relations: e.g. for a scalar field $\hat{\varphi}$, we have $[\hat{\varphi}(\vec{x},t), \frac{\partial}{\partial t}\hat{\varphi}(\vec{y},t)] = \mathrm{i}\hbar\delta^3(\vec{x}-\vec{y})$. $\delta(x)$ has the property that for any other smooth function f,

$$\int_{-1}^{1} f(y)\,\delta(y)\,\mathrm{d}y = f(0)\ , \qquad \int_{-1}^{1} f(y)\,\delta'(y)\,\mathrm{d}y = f'(0)\ ,$$

etc. The problem is that $\delta(x)$ isn't a function — no function could possibly have those properties.

One way to make sense of 'functions' like the Dirac delta and its derivatives is distribution theory. Although it was first informally used in physics, it was rigourously developed around 1950 by Laurent Schwartz, and uses the idea of test functions. See e.g. [15].

What I will describe now is an alternate approach, *algebraic* as opposed to *analytic*. These two approaches are not equivalent: you can do some things in one approach which you can't do in the other. But the algebraic approach is considerably simpler technically — no calculus or convergence to worry about — and it is remarkable how much can still be captured. This approach is the starting point for the VOA story described next section, and was first created around 1980 by Garland and Date-Kashiwara-Miwa. Keep in mind that what we are trying to capture is an operator-valued 'function' on space-time. Space-time in CFT is 2-dimensional, and so we can think of it (at least locally) as being on the complex plane $\mathbb{C}$ (more precisely, we will usually associate the space-time point (x,t) with the complex number $z = e^{t+\mathrm{i}x}$). Good introductions to the material in this section are [23,39,31].

Let W be any vector space. We are most interested in it being an infinite-dimensional space of matrices (i.e. operators on an infinite-dimensional space), but forget that for now. Define $W[[z,z^{-1}]]$ to be the set of all formal series $\sum_{n=-\infty}^{\infty} w_n z^n$, where the coefficients w_n lie in our space W. We don't ask here whether a given series converges or diverges — z is merely a formal variable. We will also be interested in $W[z,z^{-1}]$ (Laurent polynomials). We can add these formal series in the usual way, and multiply them by numbers (scalars) in the usual way.

Remember our ultimate aim here: we want to capture quantum fields. So we want our formal series to be operator-valued. The way to accomplish this is to choose W to be a vector space of operators, or matrices if you prefer. A fancy way to say this is '$W = \mathrm{End}(V)$', which means the things in W operate on vectors in V. If we take $V = \mathbb{C}^m$, then we can think of W as being the space of all $m\times m$ complex matrices. We are ultimately interested in the case $m=\infty$, but we won't lose much now by taking $m=1$, which would mean formal power series with numerical coefficients.

Because our coefficients w_n are operators, we can multiply our formal series. We define multiplication in the usual way. For example, consider $W = V = \mathbb{C}$, and take $c(z) = z^{21} - 5z^{100}$ and $d(z) = \sum_{n=-\infty}^{\infty} z^n$. Then

$$c(z)\,d(z) = \sum_{n=-\infty}^{\infty} z^{n+21} - 5\sum_{n=-\infty}^{\infty} z^{n+100} = \sum_{n=-\infty}^{\infty} z^n - 5\sum_{n=-\infty}^{\infty} z^n = -4d(z)\ .$$

So far so good. Now try to compute the square $d(z)^2$. You get infinity. So the lesson is: you can't always multiply in $W[[z,z^{-1}]]$. We'll come back to this later.

But first, look again at that first product: $c(z)\,d(z) = -4d(z)$. One thing it tells us is that *we can't always divide* (certainly $c(z)$ and -4 are two very different power series!). But there's another lesson here: if you work out a few more multiplications of this kind, what you'll find is that $f(z)\,d(z) = f(1)\,d(z)$ for any f, at least for those f for which $f(1)$ exists (e.g. any $f \in W[z,z^{-1}]$). Thus $d(z)$ is what we would call the Dirac delta $\delta(z-1)$! (You can think of it as the Fourier expansion of the Dirac delta, followed by a change of variables). Unfortunately, the standard notation here is to write it without the '-1':

$$\delta(z) := \sum_{n=-\infty}^{\infty} z^n$$

and that is the notation we will also adopt. Similarly, $\delta(az)$ and $\delta'(z)$ etc (which are the formal series defined in the obvious way) act on $W[z,z^{-1}]$ in the way one would expect: $f(z)\,\delta(az) = f(\frac{1}{a})\,\delta(az)$ and $f(z)\,\delta'(z) = f'(1)\,\delta'(z)$. So of course it makes perfect sense that we couldn't work out $d(z)^2$: we were trying to square the Dirac delta, which we know is impossible!

A similar theory can be developed for several variables z_i, with identities such as $f(z_1,z_2)\,\delta(z_1/z_2) = f(z_2,z_2)\,\delta(z_1/z_2) = f(z_1,z_1)\,\delta(z_1/z_2)$.

But we must not get too overconfident:

Paradox 1. Consider the following product:

$$\delta(z) = \big[(\sum_{n=0}^{\infty} z^n)\,(1-z)\big]\,\delta(z) = (\sum_{n=0}^{\infty} z^n)\,\big[(1-z)\,\delta(z)\big] = (\sum_{n=0}^{\infty} z^n)\,\big[0\,\delta(z)\big] = 0\ .$$

When physicists are confronted with 'paradoxes' such as this, they tend to respond by keeping them in the back of their mind, by treading with care when they are involved in a calculation which reminds them of one of the paradoxes, and otherwise trusting their instincts. Mathematicians typically over-react: they kick themselves for getting overconfident and walking head-first into a 'paradox', and then they devise some rule which will absolutely guarantee that that paradox will always be safely avoided in the future. We will follow the mathematicians' approach, and in the next few paragraphs will describe their rule for avoiding Paradox 1: to forbid certain innocent-looking products.

Remember that we are actually interested in the vector space $W = \mathrm{End}(V)$. Suppose we have infinitely many matrices $w_i \in \mathrm{End}(V)$. We will call them *summable* if for every column vector $v \in V$, only finitely many products $w_i(v) \in V$ are different from 0. In other words, only finitely many of the matrices w_i have a nonzero first column, only finitely many have a nonzero second column, ...

We will certainly have a well-defined sum $\sum_i w_i(z)$ if for each fixed n, the set $\{w_i(n)\}$ (as i varies) of matrices is summable. All other sums are forbidden. We will certainly have a well-defined product[31] $\prod_{i=1}^m w_i(z)$ if for each n, the set $\{w_1(n_1)\,w_2(n_2)\,\cdots\,w_m(n_m)\}_{\sum n_i = n}$

---

[31] m here will be finite: we permit infinite sums but only finite products.

(vary the n_i subject to $\sum_i n_i = n$) is summable. All other products are forbidden. This is reasonable because the sum of those matrix products $w_1(n_1) \cdots w_m(n_m)$ will precisely equal the nth coefficient of the product $\prod_{i=1}^m w_i(z)$.

Note that there are certainly more general ways to have a well-defined product (or sum). For example, according to our rule, we cannot even add $\sum_n 2^{-n}$! This way has the advantage of not touching the more complicated realm of convergence issues. We are doing *algebra* here, not *analysis*. The way out of Paradox 1 is that $(\sum z^n)(1-z)$ doesn't equal 1 — rather, it's a forbidden product.

An interesting consequence of the fact that we are doing algebra instead of analysis is that the product $z^{\frac{1}{2}}\delta(z)$ here does not and cannot equal $1^{\frac{1}{2}}\delta(z) = \delta(z)$ — their formal power series are very different. In hindsight this 'failing' is understandable: algebraically, it seems artificial to prefer the positive root of 1 over the negative root.

Paradox 2. Expand $\frac{1}{1-z}$ in a formal power series in z to get $\sum_{n\geq 0} z^n$. Next, expand $\frac{1}{1-z} = \frac{-z^{-1}}{1-z^{-1}}$ in a formal power series in z^{-1} to get $-\sum_{n<0} z^n$. Subtract these; we presumably should get 0, but we actually get $\delta(z)$!

The analytic explanation is that the first expression converges only for $|z| < 1$, and the second for $|z| > 1$, so it would be naive to expect their difference to be 0. We see from this 'paradox' that *it really matters in which variable we expand rational functions.* For instance, at first glance the identity

$$z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) = z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right)$$

is nonsense; it only holds if you expand the terms in positive powers of z_2, z_1, and z_0 respectively. The procedure of expanding a function in positive and negative powers of a variable and then subtracting the results, yields what are called *expansions of zero*; it is possible to show that expansions of zero will always be linear combinations of Dirac deltas $\delta(az)$ and their various derivatives $\delta^{(k)}(az)$, as we saw in Paradox 2.

2.6. INGREDIENT #3: VERTEX OPERATOR ALGEBRAS

We are now prepared to introduce the important new structure called vertex operator algebras (VOAs). They are essentially the chiral algebras of RCFTs — see [26,27] for excellent motivation of the 7 axioms below. A more detailed treatment of the basic theory of VOAs is provided by e.g. [23,39,31]. Although VOAs are natural from the CFT perspective and appear to be an important and rapidly developing area in math, their definition is not easy: Borcherds is known to have said that you either know what they are, or you don't want to know.

A VOA is a (infinite-dimensional) graded vector space $V = \oplus_{n\in\mathbb{Z}} V_n$ with infinitely many bilinear products $u *_n v$ respecting the grading (in particular $V_k *_n V_\ell \subseteq V_{k+\ell-n-1}$), which obey infinitely many constraints. 'Bilinear' means that for any $a, a', b, b' \in \mathbb{C}$ and $u, u', v, v' \in V$, $(au+a'u') *_n (bv+b'v') = ab\, u *_n v + ab'\, u *_n v' + a'b\, u' *_n v + a'b'\, u' *_n v'$ — i.e. that the products are compatible with the vector space structure of V. The subspaces V_n

must all be finite-dimensional, and they must be trivial (i.e. $V_n = \{0\}$) for all sufficiently small n (i.e. for $n \approx -\infty$). Note that we can collect all these products into one generating function: a linear map $Y : V \to (\mathrm{End}V)[[z, z^{-1}]]$. That is, to each vector $u \in V$ we associate the formal power series (called a *vertex operator*) $Y(u, z) = \sum_{n \in \mathbb{Z}} u_n z^{-n-1}$. For each u, the coefficients u_n will be functions from V to V. The idea is that the product $u *_n v$ will now be written $u_n v := u_n(v)$. The bilinearity of $*_n$ translates into two things in this new language: that $Y(\star, z)$ is linear, and that each function u_n is itself linear (i.e. they are endomorphisms).

The constraints are:

VOA 1. (*regularity*) $u_n v = 0$ for all $n > N(u, v)$;

VOA 2. (*vacuum*) there is a vector $1 \in V$ such that $Y(1, z)$ is the identity (i.e. $1_n v = \delta_{n,-1} v$);

VOA 3. (*state-field correspondence*) $Y(u, 0)1 = u$;

VOA 4. (*conformal*) there is a vector $\omega \in V$, called the *conformal vector*, such that $L_n := \omega_{n+1}$ gives us a representation of the Virasoro algebra $\mathcal{V}$, with central term $C \mapsto cI$ for some $c \in \mathbb{C}$;

VOA 5. (*translation generator*) $Y(L_{-1}u, z) = \frac{d}{dz} Y(u, z)$;

VOA 6. (*conformal weight*) $L_0 u = nu$ whenever $u \in V_n$;

VOA 7. (*locality*) $(z - w)^M [Y(u, z), Y(v, w)] = 0$ for some integer $M = M(u, v)$.

We saw the Virasoro algebra in Part 1 (see (1.2.7)). The number c in **VOA 4** is called the *central charge*, and is an important invariant of V. The peculiar-looking **VOA 7** simply says that the commutator $[Y(u, z), Y(v, w)]$ of two vertex operators will be a finite linear combination of derivatives of various orders of the Dirac delta centred at $z - w$. A recommended exercise for the reader is to show that $M = 4$ works in **VOA 7** for $u = v = \omega$. Note that in a VOA, any $Y(u, z)v$ will be a finite sum — i.e. the series $Y(u, z)$ is summable (defined last section). It is a consequence of the axioms that $1 \in V_0$ and $\omega \in V_2$: for instance, **VOA 7** says all $u_n 1 = 0$ for any $n \geq 0$, so $L_0 1 = \omega_1 1 = 0$ and hence $1 \in V_0$.

In RCFT, V would be the 'Hilbert space of states' (more carefully, V will be a dense subspace of it), and $z = e^{t+ix}$ would be a local complex coordinate on a Riemann surface. L_0 generates time translations, and so its eigenvalues (the *conformal weights*) can be identified with energy. Physically, the requirement that $V_n \to 0$ for $n \to -\infty$ corresponds to the energy of the RCFT being bounded from below. Also, $z = 0$ in **VOA 3** corresponds to the time limit $t \to -\infty$. For each state u, the vertex operator $Y(u, z)$ is a holomorphic (chiral) quantum field. The vector 1 is the vacuum $|0\rangle$, and $Y(\omega, z)$ is the stress-energy tensor T. The most important axiom, **VOA 7**, says that vertex operators commute up to a possible pole at $z = w$, and so are local quantum fields. It is equivalent to the duality axiom of many treatments of CFT. In the physics literature, there is a minor notational difference: for $u \in V_k$, $Y(u, z) = \sum u_n z^{-n-1}$ is written $\sum u_{(n)} z^{-n-k}$. (Physicists prefer this because it cleans up some formulas a little; mathematicians abhor it because it artificially prefers the 'homogeneous' vectors $u \in V_k$.)

In Segal's language (see §1.1), $Y(u, z)$ appears quite naturally. Consider the physical event of two strings combining to form a third. To first order (i.e. the tree-level Feynman diagram), this would correspond in Segal's language to a 'pair-of-pants', or a sphere with three punctures, two of which are positively oriented (corresponding to the incoming strings) and the other being negatively oriented. We can think of the sphere as the Riemann sphere $\mathbb{C} \cup \{\infty\}$; put the punctures at ∞ (outgoing) and z and 0 (the incomings). Segal's functor $\mathcal{T}$ will associate to this a z-dependent homomorphism $\varphi_z : V \times V \to V$. We write $\varphi_z(u, v) \in V$ as $Y(u, z)v$. Incidentally, the symbol 'Y' was chosen because of this 'pair-of-pants' picture (time flows from the top of the 'Y' to the bottom), as was the name 'vertex operator'.

The original axioms by Borcherds were a little more complicated and general: he didn't require $\dim(V_n) < \infty$ nor the $V_n \to 0$ condition, and he only considered L_0 and L_{-1} rather than the full Virasoro algebra. The resulting generalisation is called a *vertex algebra.*

VOA 7 can be rewritten in the form (usually called the *Jacobi identity* for the VOA)

$$z_0^{-1}\delta(\frac{z_1 - z_2}{z_0})\, Y(u, z_1)\, Y(v, z_2) - z_0^{-1}\delta(\frac{z_2 - z_1}{-z_0})\, Y(v, z_2)\, Y(u, z_1) \qquad (2.6.1)$$

$$= z_2^{-1}\delta(\frac{z_1 - z_0}{z_2})\, Y(Y(u, z_0)v, z_2)\ ,$$

where the formal series are expanded in the appropriate way. This is the embodiment of commutativity and associativity in the VOA, as we will see. To bring it into a more useful form, hit it with $t \in V$ and expand out into $z_0^\ell z_1^m z_2^n$: we obtain

$$\sum_{i\geq 0}(-1)^i \binom{\ell}{i} \left(u_{\ell+m-i} \circ v_{n+i} - (-1)^\ell v_{\ell+n-i} \circ u_{m+i}\right) = \sum_{i \geq 0} \binom{m}{i} \left(u_{\ell+i}v\right)_{m+n-i}\ , \quad (2.6.2)$$

where for any $k \in \mathbb{Z}$, $j \in \mathbb{Z}_{\geq}$, $\binom{k}{j} := \frac{k\,(k-1)\cdots(k-j+1)}{j!}$. For instance, specialising (2.6.2) to $\ell = 0$ and $m = 0$, resp., gives us

$$[u_m, v_n] = \sum_{i\geq 0} \binom{m}{i} (u_i v)_{m+n-i} \qquad (2.6.3)$$

$$(u_\ell v)_n = \sum_{i\geq 0} (-1)^i \binom{\ell}{i} \left(u_{\ell-i} \circ v_{n+i} - (-1)^\ell v_{\ell+n-i} \circ u_i\right)\ . \qquad (2.6.4)$$

Why is (2.6.1) called the Jacobi identity? Put $\ell = m = n = 0$ in (2.6.2): we get $u_0(v_0t) - v_0(u_0t) = (u_0v)_0t$. If we now formally write $[xy] := x_0y$, then this becomes $[u[vt]] - [v[ut]] = [[uv]t]$, which is one of the forms of the Lie algebra Jacobi identity (1.2.1b). Even though $[xy] \neq -[yx]$ here, this formal little trick will turn out to be quite important next section.

The simplest examples of VOAs correspond to any even positive-definite lattice Λ; for their construction see e.g. [27,39]. Physically, they correspond to a bosonic string compactified on the torus $\mathbb{R}^n/\Lambda \cong S^1 \times \cdots \times S^1$ (where n is the dimension of Λ); the central charge $c = n$. Other important examples, first constructed by Frenkel-Zhu (again see e.g. [27,39]), correspond to affine Kac-Moody algebras $X_\ell^{(1)}$ at level $k \in \mathbb{Z}_>$, and physically to WZW theories on simply-connected compact group manifolds. (We discussed affine algebras in §1.4.) These have central charge $c = \frac{k\dim(X_\ell)}{k+h^\vee}$.

In 1984 Frenkel-Lepowsky-Meurman [23] constructed the *Moonshine module* $V^\natural$. It is a VOA with $c = 24$, with $V^\natural = V_0^\natural \oplus V_1^\natural \oplus V_2^\natural \oplus \cdots$, where $V_0^\natural = \mathbb{C}1$ is 1-dimensional, $V_1^\natural = \{0\}$ is trivial, and $V_2^\natural = (\mathbb{C}\omega) \oplus (\text{Griess algebra})$ is $(1 + 196883)$-dimensional. Its automorphism group (=symmetry group) is precisely the Monster $\mathbb{M}$. Each graded piece $V_n^\natural$ is a finite-dimensional representation of $\mathbb{M}$; Borcherds proved that in fact $V^\natural$ is the McKay-Thompson infinite-dimensional representation of $\mathbb{M}$. It can be regarded as the most natural representation of $\mathbb{M}$ — it is rather surprising that important aspects of a finite group need to be studied via an infinite-dimensional representation.

$V^\natural$ has an elegant physical interpretation. First construct the bosonic string on $\mathbb{R}^{24}/\Lambda_{24}$ (recall that Λ_{24} is the Leech lattice). The resulting $c = 24$ VOA has partition function (=graded dimension) $J(\tau) + 24$, but although its graded pieces (at least for $n > 0$) have the right dimensions, they don't carry a natural representation of $\mathbb{M}$ and so can't qualify for the McKay-Thompson representation. To get $V^\natural$, orbifold this Λ_{24} VOA by the order-2 automorphism of Λ_{24} sending $\vec{x} \mapsto -\vec{x}$. $V^\natural$ thus corresponds to a holomorphic $c = 24$ RCFT, and Moonshine is related to physics. Most of Moonshine can be interpreted physically, except perhaps the genus 0 property of the McKay-Thompson series T_g.

There is a formal parallel between e.g. lattices and VOAs. For example, the Leech lattice Λ_{24} and the Moonshine module $V^\natural$ play analogous roles: Λ_{24} is the unique even lattice which (i) is self-dual, (ii) contains no norm 2 vectors, and (iii) has dimension 24; $V^\natural$ is believed to be the unique VOA which (i) possesses only one irreducible representation (namely itself), (ii) contains no conformal weight 1 elements, and (iii) has central charge $c = 24$. Analogies of these kinds are always useful as they suggest new directions to explore, and the history of math blooms with them. The battlecry 'Why invent what can be profitably copied' is not only heard in Hollywood.

We will end this section on a more speculative note. Witten (1986) said that to understand string theory conceptually, we need a new analogue of Riemannian geometry. Huang (1997) has pushed this thought a little further, saying that there is a more classical 'particle-math' and a more modern 'string-math'. According to Huang we have the real numbers (particle physics) vrs the complex numbers (string theory); Lie algebras vrs VOAs; and the representation theory of Lie algebras vrs RCFT, etc. What are the stringy analogues of calculus, ordinary differential equations, Riemannian manifolds, the Atiyah-Singer Index theorem,...? At present these are all unknown. However, Huang suggests that just as we could imagine Moonshine as a mystery which is explained in some way by RCFT, perhaps the stringy version of calculus would similarly explain the mystery of 2-dimensional gravity, stringy ODEs would explain the mystery of infinite-dimensional integrable systems, stringy Riemannian manifolds would help explain the mystery of mirror symmetry, and the stringy index theorem would help explain the elliptic genus.

2.7. INGREDIENT #4: GENERALISED KAC-MOODY ALGEBRAS

In this section we investigate Lie algebras arising from VOAs. These Lie algebras are an interesting generalisation of Kac-Moody algebras. See e.g. [5,6,38 Chapter 11.13, 29].

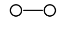
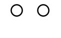
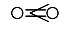
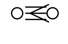
Much of Lie theory (indeed much of algebra) is developed by analogy with simple properties of integers. In §2.2 I invited you to think of a finite group as a massive generalisation of the concept of whole number. Specifically, the number n can be identified with the cyclic group $\mathbb{Z}_n$ with n elements. A divisor d of n generalises to a normal subgroup of a group. A prime number then corresponds to a simple group. Multiplying numbers corresponds to taking the semidirect product of groups (more generally, taking extensions of groups). Then we find that every group has a unique set of simple building blocks (although unlike numbers, different groups can have the same list of building blocks).

For a finite-dimensional Lie algebra, a divisor is called an ideal; a prime is called simple; and multiplying corresponds to semidirect sum. Lie algebras behave simpler than groups but not as simple as numbers, and the analogy sketched above is a reasonably satisfactory one. In particular, simple Lie algebras are important for similar reasons that simple groups are, and as mentioned in §1.2 can also be classified (with *much* less effort). A good treatment of this important classification (over $\mathbb{C}$) is provided by [36]. The proof is now reaching the state of perfection of the formulation of classical mechanics. One unobvious discovery is that the best way to capture the structure of a simple Lie algebra is by an integer matrix, called the *Cartan matrix*, or equivalently but more effectively (since most entries in the Cartan matrix are 0's) by using a graph called the *(Coxeter-)Dynkin diagram*. For instance the Dynkin diagram for A_ℓ consists of ℓ nodes connected sequentially in a line. See Figure 6 in [59].

More precisely, define a *symmetrised Cartan matrix* to be a symmetric real matrix $A = (a_{ij})_{i,j\leq\ell}$ such that $a_{ij} \leq 0$ if $i \neq j$, $a_{ii} > 0$, each $2\frac{a_{ij}}{a_{ii}} \in \mathbb{Z}$, and A is positive-definite. Examples of 2×2 symmetrised Cartan matrices are[32]

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -3 \\ -3 & 6 \end{pmatrix}$$

The Dynkin diagram corresponding to A consists of ℓ nodes; the ith and jth nodes are connected with $4a_{ij}^2/a_{ii}a_{jj}$ lines, and if $a_{ii} \neq a_{jj}$, then we put an arrow over those lines pointing to i if $a_{ii} < a_{jj}$. The Dynkin diagrams corresponding to those four Cartan matrices are respectively

 ,  ,  , 

We may without loss of generality require A to be *indecomposable*, or equivalently that the Dynkin diagram be connected. Of the 4 given above, only the second is decomposable.

To any $\ell \times \ell$ symmetrisable Cartan matrix, we can construct the corresponding Lie algebra $\mathfrak{g}$ in the following way. For each i, create 3 generators e_i, f_i, h_i (so there are a total of 3ℓ generators). The relations these generators obey are given by the following brackets: $[e_i f_j] = \delta_{ij} h_i$, $[h_i e_j] = a_{ij} e_j$, $[h_i f_j] = -a_{ij} f_j$, and for $i \neq j$ $\mathrm{ad}(e_i)^n e_j = \mathrm{ad}(f_i)^n f_j = 0$ where $n = 1 - 2\frac{a_{ij}}{a_{ii}}$. By 'ad$(e)$' here I mean the function $\mathfrak{g} \to \mathfrak{g}$ defined by $\mathrm{ad}(e)f = [ef]$. So $\mathrm{ad}(e)^2 f = [e[ef]]$, $\mathrm{ad}(e)^3 f = [e[e[ef]]]$, etc.

---

[32] Note that our Cartan matrices differ from the usual definition, in which every diagonal entry equals 2.

To get a better feeling for these relations, consider a fixed i. The generators $e = \sqrt{\frac{2}{a_{ii}}}\, e_i, f = \sqrt{\frac{2}{a_{ii}}}\, f_i, h = \frac{2}{a_{ii}}\, h_i$ obey the relations (1.2.2b). In other words, every node in the Dynkin diagram corresponds to a copy of the A_1 Lie algebra. The lines connecting these nodes tells how these ℓ copies of A_1 intertwine.

For instance consider the first Cartan matrix given above. It corresponds to the Lie algebra A_2, or $\mathrm{sl}_3(\mathbb{C})$. The two A_1 subalgebras which generate it (corresponding to the 2 nodes of the Dynkin diagram) can be chosen to be the trace-zero matrices of the form

$$\begin{pmatrix} \star & \star & 0 \\ \star & \star & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & \star & \star \\ 0 & \star & \star \end{pmatrix}.$$

It can be shown that the Lie algebra corresponding to an indecomposable symmetrised Cartan matrix will be finite-dimensional and simple, and conversely that any finite-dimensional simple Lie algebra corresponds to an indecomposable symmetrisable Cartan matrix in this way.

A confusion sometimes arises between the terms 'generators' and 'basis'. Both generators and basis vectors build up the whole algebra; the difference lies in which operations you are permitted to use. For a basis, you are only allowed to use linear combinations (i.e. addition of vectors and multiplication by numbers), while for generators you are also permitted multiplication of vectors (or the bracket, in the Lie case). 'Dimension' refers to basis, while 'rank' usually refers in some way to generators. For instance the (commutative associative) algebra of polynomials in one variable x is infinite-dimensional — any basis needs infinitely many vectors. However, the single polynomial x is enough to generate it (so we could say that its rank is 1). Although those Lie algebras have 3ℓ generators, their dimensions in general will be greater.

From the point of view of generators and relations, the step from 'finite-dimensional simple' to 'symmetrisable Kac-Moody' is rather easy: the only difference is that we drop the 'positive-definite' condition (which was responsible for finite-dimensionality). Kac-Moody (KM) algebras are also generated by (finitely many) A_1 subalgebras, and their theory is quite parallel to that of the simple algebras. Compare Figures 6 and 9 in [59].

Now, it is easy to generalise something; the challenge is to generalise it in a rich and interesting direction. One natural and appealing strategy for generalisation was followed instinctively by a grad student named Robert Moody. Moody's original motivation for developing the theory of Kac-Moody algebras was the Weyl group. If there were Lie algebras for finite Coxeter groups, he asked, why not also for the Euclidean (=affine) ones? For another example of this style of generalisation, consider the question: What is the analogue of calculus (or manifolds) over weird fields — fields (like $\mathbb{Z}_p$) for which the usual limit definitions make no sense? This question leads to the riches of algebraic geometry. Nevertheless this generalisation strategy, even in the hands of a master, will not always be successful. For instance, consider all the trouble the following metaphor has caused: my watch has a maker, so so should the Universe.

Recently Borcherds produced a further generalisation of finite-dimensional simple Lie algebras, which is rather less obvious than that of Kac-Moody algebras. It is easy to associate a Lie algebra to a matrix A, but which class of matrices will yield a deep theory?

Borcherds found such a class by holding in his hand a single algebra (the fake Monster Lie algebra, see §2.9) which acted a lot like a KM algebra, even though it had 'imaginary simple roots'.

By a *generalised symmetrised Cartan matrix* $A = (a_{ij})$ we will mean a symmetric real matrix (possibly infinite), such that $a_{ij} \leq 0$ if $i \neq j$, and if $a_{ii} > 0$ then $2\frac{a_{ij}}{a_{ii}} \in \mathbb{Z}$ for all j. By a *universal generalised Kac-Moody algebra* (universal GKM) or *universal Borcherds-Kac-Moody algebra* $\mathfrak{g}$ we mean the algebra[33] with generators e_i, f_i, h_{ij}, and with relations: $[e_i f_j] = h_{ij}$; $[h_{ij} e_k] = \delta_{ij} a_{ik} e_k$; $[h_{ij} f_k] = -\delta_{ij} a_{ik} f_k$; if $a_{ii} > 0$ and $i \neq j$ then $\mathrm{ad}(e_i)^n e_j = \mathrm{ad}(f_i)^n f_j = 0$ for $n = 1 - 2\frac{a_{ij}}{a_{ii}}$; and if $a_{ij} = 0$ then $[e_i e_j] = [f_i f_j] = 0$.

For example the Heisenberg algebra (1.2.3) corresponds to the choice $A = (0)$, while any other 1×1 $A = (a)$ corresponds to A_1. A universal GKM algebra differs from a KM algebra in that it is built up from Heisenberg algebras as well as A_1, and these subalgebras intertwine in more complicated ways. Nevertheless *much of the theory for finite-dimensional simple Lie algebras continues to find an analogue in this much more general setting* (e.g. root-space decomposition, Weyl group, character formula,...). This unexpected fact is the point of GKM algebras.

To get a feel for these algebras, let us prove a few simple results concerning the h_{ij}. Note first that, using the above relations together with the Jacobi identity, we obtain $[h_{ij} h_{k\ell}] = \delta_{ij}(a_{jk} - a_{j\ell}) h_{k\ell}$. Comparing this with $[h_{k\ell} h_{ij}] = -[h_{ij} h_{k\ell}]$, we see that bracket must always equal 0. Hence all h's pairwise commute, and $h_{ij} = 0$ unless the ith and jth columns of A are identical. An easy exercise now is to show that when $i \neq j$, h_{ij} will lie in the centre of the algebra (i.e. h_{ij} will commute with all other generators).

Although the definition of universal GKM algebra is more natural, it turns out that an equivalent form can be more useful in practice. (It's simpler to describe over $\mathbb{R}$, so in most expositions the reals are used, but alas it's far too late for us to switch loyalties now.) By a *generalised Kac-Moody algebra* (or *Borcherds-Kac-Moody algebra*) $\mathfrak{g}$, we mean a (complex!) Lie algebra which is:

- $\mathbb{Z}$-graded, i.e. $\mathfrak{g} = \oplus_i \mathfrak{g}_i$, where $[\mathfrak{g}_i \mathfrak{g}_j] \subseteq \mathfrak{g}_{i+j}$;
- $\mathfrak{g}_i$ is finite-dimensional for $i \neq 0$;
- $\mathfrak{g}$ has an antilinear involution ω (i.e. $\omega(kx+y) = k^*\omega(x) + \omega(y)$, $[\omega(x)\omega(y)] = \omega([xy])$, and $\omega \circ \omega = id$.) which maps $\mathfrak{g}_i$ to $\mathfrak{g}_{-i}$ and acts as multiplication by -1 on some basis of $\mathfrak{g}_0$;
- $\mathfrak{g}$ has an invariant symmetric bilinear form $(\star, \star)$ (i.e. $([xy], z) = (x, [yz])$ and $(y, z) = (z, y) \in \mathbb{C}$), obeying $(\omega(x), \omega(y)) = (x, y)^*$, such that $(\mathfrak{g}_i, \mathfrak{g}_j) = 0$ if $i \neq -j$;
- the Hermitian form defined by $(x|y) := -(\omega(x), y)$ is positive-definite on $\mathfrak{g}_{i\neq 0}$.

Note that for some basis x_i of $\mathfrak{g}_0$, the third condition tells us $-[x_i x_j] = [(-x_i)(-x_j)] = [x_i x_j]$, i.e. $\mathfrak{g}_0$ has a trivial bracket. It plays the role of the Cartan subalgebra $\mathfrak{h}$ in the theory.

For example, let $\mathfrak{g} = \mathrm{sl}_2(\mathbb{C})$ and recall (1.2.5). Then $\mathfrak{g}_1 = \mathbb{C}e$, $\mathfrak{g}_0 = \mathbb{C}h$, $\mathfrak{g}_{-1} = \mathbb{C}f$ is the root-space decomposition. $\omega(x) = -x^\dagger$ is the Cartan involution. $(x, y) = \mathrm{tr}(xy)$ is the Killing form.

---

[33] As with KM algebras, usually we want to extend it by some derivations; enough derivations are added so that the simple roots are linearly independent.

It turns out [5] that any universal GKM algebra is a GKM algebra, and any GKM algebra can be constructed from a unique universal GKM algebra (by quotienting out part of the centre and adding derivations), so in that sense the two structures are equivalent. This theorem is important, because it tells us that *GKM algebras are the ultimate generalisation of simple Lie algebras,* in the sense that any further generalisation will lose some basic structural ingredient.

We know simple Lie algebras (and groups) arise in both classical and quantum physics, and the affine KM algebras are important in CFT, as we saw in Part 1. GKM algebras have recently appeared in the physics literature (see Harvey-Moore) in the context of BPS states in string theory.

How do GKM algebras arise in VOAs? If we define $[xy] := x_0 y$, then as mentioned in §2.6 we get from the VOA Jacobi identity the equation $[x[yz]] - [y[xz]] = [[xy]z]$. Thus our bracket will be anti-associative if it is anti-commutative. But is it anti-commutative? It can be shown

$$u_n v = \sum_{i=0}^{\infty} \frac{1}{i!}(-1)^{i+n+1}(L_{-1})^i (v_{n+i} u) \tag{2.7.1}$$

so $u_0 v \equiv -v_0 u$ if we look at things mod $L_{-1}V$.

Since our bracket is clearly bilinear, we thus get a Lie algebra structure on $V/L_{-1}V$. Similarly, we get a symmetric bilinear product on $V/L_{-1}V$, given by $\langle u, v\rangle := u_1 v$.

We would like $\langle \star, \star\rangle$ to respect the Lie algebra structure, i.e. be $[\star\star]$-invariant. We compute from (2.6.4)

$$\langle [uv], t\rangle = -[v\langle u, t\rangle] + \langle u, [vt]\rangle \ . \tag{2.7.2}$$

Since we would like to identify $\langle \star, \star\rangle$ with the bilinear form in the GKM algebra definition, we also would like it to be number-valued (i.e. have 1-dimensional range).

There is a simple way to satisfy both of these. First, restrict attention to V_1, i.e. the conformal weight 1 vectors: $V_1 \cap (V/L_{-1}V) = V_1/L_{-1}V_0$. Then $\langle u, v\rangle \in V_0$. Assume V_0 is 1-dimensional: i.e. $V_0 = \mathbb{C}1$. Then $\langle u, v\rangle$ will equal a number times 1, so call (u, v) that number. Also, $[\langle u, t\rangle v] = (u, t)1_0 v = 0$, so $\langle \star, \star\rangle$ and hence $(\star, \star)$, will be invariant. Of course, when $V_0 = \mathbb{C}1$, $L_{-1}V_0 = \{0\}$.

$V_1/L_{-1}V_0$ is generally too large in practice to be useful; a subalgebra can be defined as follows. Let P_n be the 'primary states with conformal weight n', i.e. the $u \in V_n$ killed by L_m for all $m > 0$. Then $\mathfrak{g}(V) := P_1/L_{-1}P_0$ will be a subalgebra of $V_1/L_{-1}V_0$. Through the assignment $u \mapsto u_0$, $\mathfrak{g}(V)$ acts on V and this action commutes with that of L_i. This association of a Lie algebra to a VOA is due to Borcherds (1986).

Similar arguments show that when V_0 is 1-dimensional and V_1 is 0-dimensional, then V_2 will necessarily be a commutative nonassociative algebra with product $u \times v := u_1 v \in V_2$ and identity element $\frac{1}{2}\omega$ (*proof*: $\omega \times u = L_0 u = 2u$). Now, those conditions on V_0, V_1 are satisfied by the Moonshine module $V^\natural$. We find that $V_2^\natural$ is none other than the Griess algebra extended by an identity element.

2.8. INGREDIENT #5: DENOMINATOR IDENTITIES

In §1.3, we discussed the representation theory of Lie algebras. An important invariant of a representation is its *character*. Simple Lie algebras possess a very useful formula for their characters, due to Weyl:

$$\mathrm{ch}_\lambda(z) := \sum_\mu \dim(L_\lambda(\mu))\, e^{\mu\cdot z} = e^{-\rho\cdot z}\,\frac{\sum_{w\in W} \pm e^{w(\lambda+\rho)\cdot z}}{\prod_{\alpha\in\Delta_+}(1-e^{-\alpha\cdot z})}\ , \tag{2.8.1}$$

where W is the Weyl group, Δ_+ the positive roots, and where $\oplus_\mu L_\lambda(\mu)$ is the weight-space decomposition of L_λ — i.e. the simultaneous eigenspaces of the h_i. Here z belongs to the Cartan subalgebra $\mathfrak{h}$; the character is complex-valued. Analogous statements hold for all GKM algebras.

It is rare indeed when a trivial special case of a theorem or formula produces something interesting. But that is what happens here. Consider the trivial representation: i.e. $x \mapsto 0$ for all $x \in X_\ell$. Then the character is identically 1, by definition: $ch_0 \equiv 1$. Thus the character formula tells us that a certain alternating sum over a Weyl group, equals a certain product over positive roots. These formulas, called *denominator formulas*, are nontrivial even in the finite-dimensional cases.

Consider for instance the smallest simple algebra, A_1. Here the identity indeed is too trivial: it reads $e^{z/2} - e^{-z/2} = e^{z/2}(1-e^{-z})$. For A_2 we get a sum of 6 terms equalling a product of 3 terms, and the complexity continues to rise from there.

Around 1970 Macdonald tried to generalise these finite denominator identities to infinite identities, corresponding to the extended Dynkin diagrams. These were later reinterpreted by Kac, Moody and others as denominator identities for affine nontwisted KM algebras. The simplest one was known classically as the Jacobi triple product identity:

$$\sum_{n=-\infty}^{\infty}(-1)^n x^{n^2} y^n = \prod_{m=1}^{\infty}(1-x^{2m})(1-x^{2m-1}y)(1-x^{2m-1}y^{-1})$$

We now know it to be the denominator identity for the simplest infinite-dimensional KM algebra, $A_1^{(1)}$.

Freeman Dyson is famous for his work in quantum field theory, but he started as an undergraduate in number theory and still enjoys it as a hobby. Dyson [20] found a curious formula for the Ramanujan τ-function, which can be defined by the generating function $\sum_{n=1}^{\infty}\tau(n)x^n = \eta^{24}(x) := x\prod_{m=1}^{\infty}(1-x^m)^{24}$. Dyson found the remarkable formula

$$\tau(n) = \sum \frac{(a-b)(a-c)(a-d)(a-e)(b-c)(b-d)(b-e)(c-d)(c-e)(d-e)}{1!\ 2!\ 3!\ 4!}$$

where the sum is over all 5-tuples $(a,b,c,d,e) \equiv (1,2,3,4,5) \pmod 5$ obeying $a+b+c+d+e=0$ and $a^2+b^2+c^2+d^2+e^2=10n$. Using this, an analogous formula can be found for η^{24}. Dyson knew that similar-looking formulas were also known for η^d for the values $d = 3, 8, 10, 14, 15, 21, 24, 26, 28, 35, 36, \ldots$.

What was ironic was that Dyson found that formula at the same time that Macdonald was finding the Macdonald identities. Both were at Princeton then, and would often chat a little when they bumped into each other after dropping off their daughters at school. But they never discussed work. Dyson didn't realise that his strange list of numbers has a simple interpretation: they are precisely the dimensions of the simple Lie algebras! $3 = \dim(\mathrm{A}_1)$, $8 = \dim(A_2)$, $10 = \dim(C_2)$, $14 = \dim(G_2)$, etc. In fact these formulas for η^d are none other than (specialisations of) the Macdonald identities. For example, Dyson's formula is the denominator formula for $A_4^{(1)}$ ($24 = \dim(A_4)$). If they had spoken, they would probably have anticipated the affine denominator identity interpretation.

One curiousity apparently still has no algebraic interpretation: No simple Lie algebra has dimension 26, so the formula for η^{26} can't correspond to any Macdonald identity.

Macdonald didn't close the book on denominator identities. More recently Kac and Wakimoto [40] have used denominator identities for Lie superalgebras to obtain nice formulas for various generator functions involving sums of squares, sums of triangular numbers (triangular numbers are numbers of the form $\frac{1}{2}k(k+1)$), etc. For instance, the number of ways n can be written as a sum of 16 triangular numbers is

$$\frac{1}{3 \cdot 4^3} \sum ab\,(a^2 - b^2)^2$$

where the sum is over all odd positive integers a, b, r, s obeying $ar + bs = 2n + 4$ and $a > b$.

Another example of denominator identities is Borcherds' use of them in proving the Moonshine conjectures. In particular this motivated his introduction of the GKM algebras. The denominator identities for other GKM algebras were used by Borcherds to obtain results on the moduli spaces of e.g. families of K3 surfaces. They are also often turned-around now and used for learning about the positive roots in a given GKM.

2.9. Proof of the Moonshine conjectures

The main Conway-Norton conjecture was proved almost immediately. Thompson showed that if $g \mapsto a_n(g)$ is a character for all $n \leq 1200$, then it will be for all n. He also showed that if certain congruence conditions hold for a certain number of $a_n(g)$ (all with $n \leq 100$), then all $g \mapsto a_n(g)$ will be *virtual* characters (i.e. a linear combination over $\mathbb{Z}$ of irreducible characters of $\mathbb{M}$; only if all coefficients are nonnegative will it be a true character). Atkin-Fong-Smith [54] used that to prove on a computer that indeed all were virtual characters. But their work didn't say anything about the underlying (possibly virtual) representation V. The real challenge was to construct (preferably in a natural manner) the representation which works. Frenkel-Lepowsky-Meurman [23] constructed a candidate for it (the Moonshine module $V^\natural$); it was Borcherds who finally proved $V^\natural$ obeyed the Conway-Norton conjecture. A good overview of Borcherds' work on Moonshine is provided in [32].

We want to show that the McKay-Thompson series $T_g(\tau) := q^{-1}\mathrm{tr}_{V^\natural}(gq^{L_0})$ of (2.1.4) equals the Hauptmodul $J_g(\tau)$ in (2.4.1) (the 'fudge factor' $q^{-1} = q^{-c/24}$ is familiar to e.g. KM algebras and CFT and was discussed at the end of §1.2). Borcherds' strategy was to bring in Lie theory and to use the corresponding denominator identity to provide useful combinatorial data. The first guess for this 'Monster Lie algebra' was the Kac-Moody algebra whose Dynkin diagram is essentially the Leech lattice (i.e. a node for each vector in Λ_{24}, and 2 nodes are connected by a number of edges depending on the value of a certain dot product). It was eventually discarded because some of the critical data (namely positive root multiplicities) needed in order to write down its denominator identity were too complicated. But looking at that failed attempt led Borcherds to a second candidate, now called the *fake Monster Lie algebra* $\mathfrak{g}'_{\mathbb{M}}$. In order to *construct* it, he developed the theory of VOAs; and in order to *understand* it, he developed the theory of GKM algebras. We will define it shortly. $\mathfrak{g}'_{\mathbb{M}}$ also turned out to be inadequate for proving the Moonshine conjectures; however it directly led him to the GKM algebra now called the true *Monster Lie algebra* $\mathfrak{g}_{\mathbb{M}}$. And that directly led to the proof of Moonshine.

Step 1: Construct $\mathfrak{g}_{\mathbb{M}}$ from $V^\natural = V_0^\natural \oplus V_1^\natural \oplus \cdots$. For later convenience, reparametrise these subspaces $V^i := V^\natural_{i+1}$. Recall the even indefinite lattice $II_{1,1}$ defined in (1.6.1). Of course the direct choice $\mathfrak{g}(V^\natural)$ is 0-dimensional because $V_1^\natural$ is trivial, so we must modify $V^\natural$ first.

The Monster Lie algebra is (essentially) defined to be the Lie algebra $\mathfrak{g}(V^\natural \otimes V_{II_{1,1}})$ associated to the vertex algebra $V^\natural \otimes V_{II_{1,1}}$ (strictly speaking, more of $\mathfrak{g}$ is quotiented away). By contrast, the fake Monster is the Lie algebra associated to the vertex algebra $V_{\Lambda_{24}} \otimes V_{II_{1,1}} \cong V_{II_{25,1}}$. Both of these are vertex algebras as opposed to VOAs, because of the presence of the indefinite lattices, but this isn't important here. $\mathfrak{g}_{\mathbb{M}}$ inherits a $II_{1,1}$-grading from $V_{II_{1,1}}$: the piece of grading (m,n) is isomorphic (as a vector space) to V^{mn}, if $(m,n) \neq (0,0)$; the (0,0) piece is isomorphic to $\mathbb{R}^2$. Borcherds uses the No-Ghost Theorem of string theory to show that the homogeneous pieces of $\mathfrak{g}_{\mathbb{M}}$ are those of $V^\natural$.

Both $\mathfrak{g}_{\mathbb{M}}$ and $\mathfrak{g}'_{\mathbb{M}}$ are GKM algebras; for instance the $\mathbb{Z}$-grading of $\mathfrak{g}_{\mathbb{M}}$ is given by $(\mathfrak{g}_{\mathbb{M}})_k = \oplus_{m+n=k} V^{mn}$ for $k \neq 0$, while the 0-part is $V^{-1} \oplus V^{-1}$. Although $\mathfrak{g}'_{\mathbb{M}}$ is not used in the proof of the Monstrous Moonshine conjectures, it is related to some kind of Moonshine for the finite simple group .1, which is 'half' of the automorphism group .0 of the Leech lattice Λ_{24}. $\mathfrak{g}_{\mathbb{M}}$ corresponds to the Cartan matrix

$$\begin{pmatrix} B_{-1,-1} & B_{-1,1} & B_{-1,2} & \cdots \\ B_{1,-1} & B_{1,1} & B_{1,2} & \cdots \\ B_{2,-1} & B_{2,1} & B_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} ,$$

where $B_{i,j}$ for $i,j \in \{-1,1,2,3,\ldots\}$ is the $a_i \times a_j$ block with fixed entry $-i-j$ (the a_i as usual are the coefficients $\sum_n a_n q^n$ of $j - 744$).

2nd step: Compute the denominator identity of $\mathfrak{g}_{\mathbb{M}}$: we get

$$p^{-1} \prod_{\substack{m>0 \\ n\in\mathbb{Z}}} (1-p^m q^n)^{a_{mn}} = j(z) - j(\tau) \tag{2.9.1}$$

where $p = e^{2\pi \mathrm{i} z}$. The result are various formulas involving the coefficients a_i, for instance $a_4 = a_3 + (a_1^2 - a_1)/2$. It turns out to be possible to 'twist' (2.9.1) by each $g \in \mathbb{M}$, obtaining

$$p^{-1} \exp[-\sum_{k>0} \sum_{\substack{m>0 \\ n\in\mathbb{Z}}} a_{mn}(g^k) \frac{p^{mk} q^{nk}}{k}] = T_g(z) - T_g(\tau) \ . \tag{2.9.2}$$

This looks a lot more complicated, but you can glimpse the Taylor expansion of $\ln(1-p^m q^n)$ there and in fact for $g = id$ (2.9.2) reduces to (2.9.1). This formula gives more generally identities like $a_4(g) = a_2(g) + (a_1(g)^2 - a_1(g^2))/2$, where $T_g(\tau) = \sum_i a_i(g)q^i$. These formulas involving the McKay-Thompson coefficients are equivalent to the replication formulae conjectured in §2.4.

3rd step: It was known earlier that all of the Hauptmoduls also obey those replication formula, and that anything obeying them will lie in a finite-dimensional manifold which we'll call R. In particular, if $B(q) = q^{-1} + \sum_{n>0} b_n q^n$ and $C(q) = q^{-1} + \sum_{n>0} c_n q^n$ both lie in R, and $b_n = c_n$ for $n \leq 23$, then $B(q) = C(q)$. In fact, it turns out that if we verify for each conjugacy class $[g]$ of $\mathbb{M}$ that the first, second, third, fourth and sixth coefficients of the McKay-Thompson series T_g and the corresponding Hauptmodul J_g agree, then $T_g = J_g$, and we are done.

That is precisely what Borcherds then did: he compared finitely many coefficients, and as they all equalled what they should, this concluded the proof of Monstrous Moonshine!

However there was a disappointing side to his proof. While no one disputed its *logical* validity, it did seem to possess a disappointing *conceptual gap*. In particular, the Moonshine conjectures were made in the hope that proving them would help explain what the Monster had to do with the j-function and the other Hauptmoduls. A good proof says much more than 'True' or 'False'. The case-by-case verification occurred at the critical point where the McKay-Thompson series were being compared directly to the Hauptmoduls. The proof showed that indeed the Moonshine module establishes some sort of relation between T_g and J_g (namely, they must lie in the same finite-dimensional space), but why couldn't it be just a happy meaningless accident that they be equal? Of course we believe it's more than merely an accident, so our proof should reflect this: we want a more conceptual explanation.

This conceptual gap has since been filled [17] — i.e. the case-by-case verification has been replaced with a general theorem. It turns out that something obeying the replicable formulas will also obey something called *modular equations*. A modular equation for a function f is a polynomial identity obeyed by $f(x)$ and $f(nx)$. The simplest examples come from the exponential and cosine functions: note that for any $n > 0$, $\exp(nx) = (\exp(x))^n$ and $\cos(nx) = T_n(\cos(x))$ where T_n is a Tchebychev polynomial. A more interesting example of a modular equation is obeyed by $J(\tau) = j(\tau) - 744$: put $X = J(\tau)$ and $Y = J(2\tau)$, then

$$\begin{aligned}(X^2 - Y)(Y^2 - X) = {} & 393768\,(X^2 + Y^2) + 42987520\,XY + 40491318744\,(X + Y) \\ & - 120981708338256\ .\end{aligned}$$

Finding modular equations (for various elliptic functions) was a passion of the great mathematician Ramanujan. His notebooks are filled with them. See e.g. [7] for an application of Ramanujan's modular equations to computing the first billion or so digits of π. Many modular equations are also studied in [10]. For more of their applications, see e.g. [16]. It can be shown that the only functions $f(\tau) = q^{-1} + a_1 q + \cdots$ which obey modular equations for all n, are $J(\tau)$ and the 'modular fictions' q^{-1} and $q^{-1} \pm q$ (which are essentially exp, cos, and sin).

It was proved in [17] that, roughly speaking, a function $B(\tau)=q^{-1}+\sum_{n>0}b_nq^n$ which obeys enough modular equations, will either be of the form $B(\tau)=q^{-1}+b_1q$, or will necessarily be a Hauptmodul for a modular group containing some $\Gamma_0(N)$. The converse is also true: for instance, a modular equation for the Hauptmodul J_{25} of $\Gamma_0(25)$ given in (2.3.5) is

$$(X^2-Y)(Y^2-X)=-2\,(X^2+Y^2)+4\,(X+Y)-4\ ,$$

where $X=J_{25}(\tau)$ and $Y=J_{25}(2\tau)$. To eliminate the conceptual gap, this result should then replace step 3. Steps 1 and 2 are still required, however.

This conceptual gap should not take away from what was a remarkable accomplishment by Borcherds: not only the proof of the Monstrous Moonshine conjectures, but also the definition of two new and important algebraic structures. I hope the preceding sections give the reader some indication of why Borcherds was awarded one of the 1998 Fields Medals.

Another approach to the Hauptmodul property is by Tuite [57], who related it to the (conjectured) uniqueness of $V^\natural$. Norton has suggested that the reason $\mathbb{M}$ is associated to *genus-zero* modular functions could be what he calls its '6-transposition' property [47].

So has Moonshine been explained? According to Conway, McKay, and many others, it hasn't. They consider VOAs in general, and $V^\natural$ in particular, to be too complicated to be God-given. The progress, though impressive, has broadened not lessened the fundamental mystery, they would argue.

For what it's worth, I don't completely agree. Explaining away a mystery is a little like grasping a bar of soap in a bathtub, or quenching a child's curiousity. Only extreme measures like pulling the plug, or enrollment in school, ever really work. True progress means displacing the mystery, usually from the particular to the general. Why is the sky blue? Because of how light scatters in gases. Why are Hauptmoduls attached to each $g\in\mathbb{M}$? Because of $V^\natural$. Mystery exists wherever we can ask 'why' — like beauty it's in the beholder's eye.

Moonshine is now 'leaving the nest'. We are entering a consolation phase, tidying up, generalising, simplifying, clarifying, working out more examples. Important and interesting discoveries will be made in the next few years, and yes there still is mystery, but no longer does a Moonshiner feel like an illicit distiller: Moonshine is now a day-job!

Acknowledgments. I warmly thank the Feza Gursey Institute in Istanbul, and in particular Teoman Turgut, for their invitation to the workshop and hospitality during my month-long stay. These notes are based on 16 lectures I gave there in Summer 1998. I've also benefitted from numerous conversations with Y. Billig, A. Coste, C. Cummins, M. Gaberdiel, J. McKay, M. Tuite, and M. Walton — Mark Walton in particular made a very careful reading of the manuscript (and hence must share partial blame for any errors still remaining). My appreciation as well goes to J.-B. Zuber and P. Ruelle for sharing with me their personal stories behind the discoveries of, respectively, A-D-E in $A_1^{(1)}$ and Fermat in $A_2^{(1)}$. The research was supported in part by NSERC.

References

1. V. I. Arnold, *Catastrophe Theory*, 2nd edn. (Springer, Berlin, 1997);
   M. Hazewinkel, W. Hesselink, D. Siersma, and F.D. Veldkamp, *Nieuw Arch. Wisk.* **25** (1977) 257;
   P. Slodowy, in: Lecture Notes in Math 1008, J. Dolgachev (ed.) (Springer, Berlin, 1983).
2. M. Bauer, A. Coste, C. Itzykson, and P. Ruelle, *J. Geom. Phys.* **22** (1997) 134.
3. D. Bernard, *Nucl. Phys.* **B288** (1987) 628.
4. J. Böckenhauer and D. E. Evans, *Commun. Math. Phys.* **200** (1999) 57; "Modular invariants, graphs and α-induction for nets of subfactors III", hep-th/9812110.
5. R. E. Borcherds, *Invent. math.* **109** (1992) 405.
6. R. E. Borcherds, "What is moonshine?", math.QA/9809110.
7. J. M. Borwein, P. B. Borwein and D. H. Bailey, *Amer. Math Monthly* **96** (1989) 201.
8. A. Cappelli, C. Itzykson, and J.-B. Zuber, *Commun. Math. Phys.* **113** (1987) 1.
9. R. Carter, G. Segal and I. M. Macdonald, *Lectures on Lie Groups and Lie algebras* (Cambridge University Press, Cambridge, 1995).
10. A. Cayley, *An Elementary Treatise on Elliptic Functions*, 2nd edn (Dover, New York, 1961).
11. J. H. Conway, *Math. Intelligencer* **2** (1980) 165.
12. J. H. Conway and S. P. Norton, *Bull. London Math. Soc.* **11** (1979) 308.
13. J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, 3rd edn (Springer, Berlin, 1999).
14. A. Coste and T. Gannon, *Phys. Lett.* **B323** (1994) 316.
15. R. Courant and D. Hilbert, *Methods of Mathematical Physics II* (Wiley, New York, 1989).
16. D. Cox, *Primes of the form $x^2 + ny^2$*, (Wiley, New York, 1989).
17. C. J. Cummins and T. Gannon, *Invent. math.* **129** (1997) 413.
18. P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory* (Springer, New York, 1996).
19. L. Dixon, P. Ginsparg, and J. Harvey, *Commun. Math. Phys.* **119** (1988) 221.
20. F. Dyson, *Bull. Amer. Math. Soc.* **78** (1972) 635.
21. F. J. Dyson, *Math. Intelligencer* **5** (1983) 47.
22. D.E. Evans and Y. Kawahigashi, *Quantum symmetries on operator algebras* (Oxford University Press, Oxford, 1998).
23. I. Frenkel, J. Lepowsky, and A. Meurman, *Vertex operator algebras and the Monster* (Academic Press, San Diego, 1988).
24. J. Fuchs and C. Schweigert, *Symmetries, Lie algebras, and representations* (Cambridge University Press, Cambridge, 1997).
25. W. Fulton and J. Harris, *Representation Theory: A first course* (Springer, New York, 1996).
26. M. R. Gaberdiel and P. Goddard, "Axiomatic conformal field theory", hep-th/9810019.
27. M. R. Gaberdiel and P. Goddard, "An introduction to meromorphic conformal field theory and its representations", lecture notes in this volume.

28. D. Gaitsgory, "Notes on two dimensional conformal field theory and string theory", math.AG/9811061.
29. T. Gannon, "The Cappelli-Itzykson-Zuber A-D-E classification", math.QA/9902064.
30. K. Gawedzki, "Conformal field theory: a case study", lecture notes in this volume.
31. R. W. Gebert, *Internat. J. Mod. Phys.* **A8** (1993) 5441.
32. P. Goddard, "The work of R.E. Borcherds", math.QA/9808136.
33. D. Gorenstein, *Finite Simple Groups* (Plenum, New York, 1982).
34. D. Gorenstein, R. Lyons, and R. Solomon, *The Classification of the Finite Simple Groups* (AMS, Providence, 1994).
35. A. Hanany and Y.-H. He, "Non-abelian finite gauge theories", hep-th/9811183.
36. J. E. Humphreys, *Introduction to Lie algebras and representation theory* (Springer, New York, 1994).
37. V. G. Kac, In: Lecture Notes in Math 848 (Springer, New York, 1981).
38. V. G. Kac, *Infinite Dimensional Lie algebras*, 3rd edn (Cambridge University Press, Cambridge, 1990).
39. V. G. Kac, *Vertex Algebras for Beginners*, 2nd edn (AMS, Providence, 1998).
40. V. G. Kac and M. Wakimoto, In: *Lie Theory and Geometry in Honor of Bertram Kostant*, Progress in Math. **123** (Birkhäuser, Boston, 1994).
41. S. Kass, R. V. Moody, J. Patera and R. Slansky, *Affine Lie algebras, weight multiplicities, and branching rules*, Vol. 1 (University of California Press, Berkeley, 1990).
42. S. Lang, *Elliptic Functions*, 2nd edn (New York, Springer, 1997).
43. F. W. Lawvere and S. H. Schanuel, *Conceptual Mathematics: A first introduction to Categories* (Cambridge University Press, Cambridge, 1997).
44. H. Minc, *Nonnegative matrices* (Wiley, New York, 1988).
45. E.J. Mlawer, S.G. Naculich, H.A. Riggs, and H. J. Schnitzer, *Nucl. Phys.* **B352** (1991) 863.
46. W. Nahm, *Commun. Math. Phys.* **118** (1988) 171.
47. S. P. Norton, In: *Proc. Symp. Pure Math.* **47** (1987) 208.
48. A. Ocneanu, "Paths on Coxeter diagrams: From Platonic solids and singularities to minimal models and subfactors" (Lectures given at Fields Institute (1995), notes recorded by S. Goto).
49. L. Queen, *Math. of Comput.* **37** (1981) 547.
50. A.N. Schellekens and S. Yankielowicz, *Nucl. Phys.* **B327** (1989) 673.
51. M. Schottenloher, *A Mathematical Introduction to Conformal Field Theory* (Springer, Berlin, 1997).
52. G. Segal, In: *IXth Proc. Int. Congress Math. Phys.* (Hilger, 1989).
53. S. Singh, *Fermat's Enigma* (Penguin Books, London, 1997).
54. S. D. Smith, In: *Finite Groups – Coming of Age*, Contemp. Math. **193** (AMS, Providence, 1996).
55. W. Thurston, *Three-dimensional geometry and topology*, vol. 1 (Princeton, 1997).
56. L. Toti Rigatelli, *Evariste Galois* (Birkhäuser, Basel, 1996).
57. M. Tuite, *Commun. Math. Phys.* **166** (1995) 495.
58. V. G. Turaev, *Quantum invariants of knots and 3-manifolds* (de Gruyter, Berlin, 1994).

59. M. A. Walton, “Affine Kac-Moody algebras and the Wess-Zumino-Witten model”, lecture notes in this volume.
60. M. Waldschmidt, P. Moussa, J.-M. Luck, and C. Itzykson (ed.), *From Number Theory to Physics* (Berlin, Springer, 1992).
61. D. Zagier, *Amer. Math. Monthly* **97** (1990) 144.
62. J.-B. Zuber, *Commun. Math. Phys.* **179** (1996) 265.
63. J.-B. Zuber, private communication, February 1999.

arXiv:math/9906167v2 [math.QA] 10 Aug 1999

August, 1999

Monstrous Moonshine and the Classification of CFT

(16 lectures given in Istanbul, August 1998)

Terry Gannon

Department of Mathematical Sciences, University of Alberta,
Edmonton, Alberta, Canada, T6G 2G1

e-mail: tgannon@math.ualberta.ca

Abstract

In these notes we give an introduction both to Monstrous Moonshine and to the classification of rational conformal field theories, using this as an excuse to explore several related structures and go on a little tour of modern math. We will discuss Lie algebras, modular functions, the finite simple group classification, vertex operator algebras, Fermat's Last Theorem, category theory, (generalised) Kac-Moody algebras, denominator identities, the A-D-E meta-pattern, representations of affine algebras, Galois theory, etc. This work is informal and pedagogical, and aimed mostly at grad students in math or math phys, but I hope that many interested nonexperts will find something of value here — like any good Walt Disney movie I try not to completely ignore the 'grown-ups'. My emphasis is on ideas and motivations, so these notes are intended to complement other papers and books where this material is presented with more technical detail. The level of difficulty varies significantly from topic to topic. The two parts — in fact any of the sections — can be read independently.

Table of Contents

Part 1. The classification of conformal field theories

1.1. Informal motivation 1
1.2. Lie algebras 4
1.3. Representations of finite-dimensional simple Lie algebras 10
1.4. Affine algebras and the Kac-Peterson matrices 13
1.5. The classification of physical invariants 17
1.6. The A-D-E meta-pattern 19
1.7. Simple-currents and charge-conjugation 24
1.8. Galois theory 28
1.9. The modern approach to classifying physical invariants 31

Part 2. Monstrous Moonshine

2.1. Introduction 35
2.2. Ingredient #1: Finite simple groups and the Monster 38
2.3. Ingredient #2: Modular functions and Hauptmoduls 40
2.4. The Monstrous Moonshine conjectures 46
2.5. Formal power series 48
2.6. Ingredient #3: Vertex operator algebras 51
2.7. Ingredient #4: Generalised Kac-Moody algebras 54
2.8. Ingredient #5: Denominator identities 59
2.9. Proof of the Moonshine conjectures 60

References 64

Glossary

24	§1.6
A_ℓ	$\mathrm{sl}_{\ell+1}(\mathbb{C})$
affine algebra	§1.4
algebra	§1.2
category theory	§1.1
central extension	§§1.2, 1.4
CFT	conformal field theory
character	§§1.3, 1.4
chiral algebra	§1.1
Fermat's Last Theorem	§2.3
field	§1.2
finite simple group	§2.2
genus	§§1.1, 2.3
group	§§1.3, 2.2
Hauptmodul	§2.3
highest-weight	§§1.3, 1.4
Kac-Moody algebra	§2.7
lattice	§1.6
Leech lattice Λ_{24}	§2.4
Lie algebra	§1.2
Lie group	§1.2
$\mathbb{M}$	the Monster finite simple group
manifold	§1.2
modular function	§2.3
modular group	$\mathrm{PSL}_2(\mathbb{Z})$ or $\mathrm{SL}_2(\mathbb{Z})$
P_+^k	§1.4
partition function	§§1.1, 1.2, 1.5
Perron-Frobenius	§1.7
physical invariant	RCFT partition function, §1.5
QFT	quantum field theory
RCFT	rational conformal field theory
representation	§1.3
Riemann surface	§§1.1, 2.3
S, T	Kac-Peterson matrices, (1.4.2)
subfactor theory	§1.9
topological	§§1.1, 2.3
torus	§2.3
$V^\natural$	the Moonshine module, §§2.4, 2.6
Verlinde's formula	(1.4.4)
Virasoro algebra	(1.2.7)
VOA	vertex operator algebra
Weyl group	§1.3
WZW model	§1.1

Part 1. The classification of conformal field theories

1.1. INFORMAL MOTIVATION

In this section we will sketch a very informal and 'hand-wavy' motivation to what we shall call the classification problem for rational conformal field theory (RCFT). Much of this material is more carefully treated in e.g. [18].

A CFT is a quantum field theory (QFT), usually with a two-dimensional space-time, whose symmetries include the conformal transformations. There are different approaches to CFT — for one of these see [26,27]. Another formulation which has been deeply influential is due to Graeme Segal [52]. It is motivated by string theory and is phrased in an important mathematical language called *category theory*.

A *category* consists of two types of things. One are called *objects*, and the other are called *arrows* (or *morphisms*). An arrow, written $f : A \rightarrow B$, has an initial and a final object (A and B respectively). Arrows f, g can be composed to yield a new arrow $f \circ g$, if the final object of g equals the initial object of f. Maps between categories are called *functors* if they take the objects (resp., arrows) of one to the objects (resp., arrows) of the other, and preserve composition.

The only difficulty people can have in understanding categories is in realising that there is no real content to them. It's just a language, highly abstract like the more familiar set theory, but in many contexts (a great example is the theory of knot invariants [58]) one which is both natural and suggestive. It tries to deflect some of our instinctive infatuation with objects (nouns), to the mathematically more fruitful one with structure-preserving maps between objects (verbs). A gentle introduction to the mathematics and philosophy of categories is [43]; we'll give a taste of this shortly.

The standard example of a category is called **Set**, where the 'objects' are sets, and the arrows from A to B are functions $A \rightarrow B$. A related example that Segal uses is **Vect**, where the objects are complex vector spaces and the arrows are linear maps. A rather trivial example of a functor $\mathcal{F} : \mathbf{Vect} \rightarrow \mathbf{Set}$ sends a vector space V to its underlying set, also called V — i.e. $\mathcal{F}$ simply 'forgets' the vector space structure on V and ignores the fact that the arrows f in **Vect** are linear. The other category Segal uses he calls **C**; its objects are disjoint unions of (parametrised) circles S^1, and the arrows are (conformal equivalence classes of) *cobordisms*, i.e. (Riemann) surfaces whose boundaries are those circles. Composition of arrows in **C** is accomplished by gluing the surfaces along the appropriate boundary circles.

Consider the usual definition of a one-to-one function: $f(x) = f(y)$ only when $x = y$. Category theory replaces this with the following. The arrow $f : A \rightarrow B$ is called 'monic' if for any arrow $g : C \rightarrow B$, there exists a unique arrow $h : C \rightarrow A$ such that $f \circ h = g$. So it's a sort of factorisation property. You can easily verify that in **Set** the notions of 'one-to-one' and 'monic' coincide. What does this redefinition gain us? It certainly doesn't seem any simpler. But it does change the focus from the *argument* of f, to the *global* functional behaviour of f, and a change of perspective can never be bad. And it allows us to transport the idea of one-to-one-ness to arbitrary categories. For instance, in the Riemann surface category **C**, the 'one-to-one functions' are the genus-0 cobordisms.

Or consider the notion of *product*. In category theory, we say that the triple (P, a, b) is a product of objects A, B if $a : P \to A$ and $b : P \to B$ are arrows, and if for any $f : C \to A$, $g : C \to B$, there is a unique arrow $h : C \to P$ such that $f = a \circ h$ and $g = b \circ h$. This notion unifies several constructions (each of which is the 'product' in an appropriately chosen category): Cartesian product of sets; intersection of sets; multiplication of numbers; the logical operator 'and'; direct product; infimum in a partially ordered set; etc. *Sum* can be defined similarly, unifying the constructions of disjoint union, 'or', addition, tensor product, direct sum, supremum, etc.

This generality of course comes with a price: it can wash away all of the endearing special features of a favourite theory or structure. There certainly are contexts where e.g. all human beings should be thought of as equal, but there are other contexts where the given human is none other than your mother and must be treated as such. It turns out that category theory provides a beautiful framework for understanding topological invariants such as the Jones-Reshetikhin-Turaev-Witten knot invariants (see e.g. [58]). And it seems to be a natural language for formulating CFT axiomatically, as we'll now see.

According to Segal, a CFT is a functor $\mathcal{T}$ from $\mathbf{C}$ to $\mathbf{Vect}$, which obeys various properties. The picture comes from string theory: the fundamental object is a 'vibrating' loop; a state is given by a collection of these loops; each classical Feynman path from the initial to the final states is a world-sheet, i.e. a surface Σ whose boundary is all the loops in the initial and final states. QFT assigns a complex number I (the action) to each of these world-sheets, and the quantum amplitude, written $\langle \text{final} | \text{initial} \rangle$, will then be the integral over all worldsheets of $e^{\mathrm{i}I/\hbar}$. (The quantum amplitude is how the theory makes contact with experiment, as it tells us the probability of the given process $|\text{initial}\rangle \mapsto |\text{final}\rangle$ happening.) This is what Segal is trying to capture formally. The vector spaces in $\mathbf{Vect}$ come in in order to handle uniformly and simultaneously the various 'vibrations' of the strings. In particular there is one basic vector space H (a Hilbert space of quantum states), and the functor will take n copies of S^1 to $H^n = H \otimes \cdots \otimes H$.

The simplest interesting example here is the 'tree-level creation of a string from the vacuum'. In this case the world-sheet looks like a bowl, i.e. *topologically* is a disk D (if we imagine the bowl to be made of rubber, we could grab its rim and stretch it down flat onto the table, so we say a bowl and a disk are topologically equivalent — see also §2.3). Segal's functor gives us a linear map $\mathcal{T}(D) : \mathbb{C} \to H$ (H^0 is just $\mathbb{C}$), which we can think of equivalently as the assignment of the vector $\mathcal{T}(D)(1)$ in H to D. In the case of the standard unit disk (i.e. where the parametrisation of the boundary S^1 is simply $\theta \mapsto e^{2\pi \mathrm{i}\theta}$), this vector is cal0led the *vacuum state* $\Omega = |0\rangle$.

For another example, consider the 'vacuum-to-vacuum expectation value'. The initial and final states (objects) here are both the empty set, so the world-sheets (arrows) are closed Riemann surfaces. As usual in QFT, we can organise these by how many internal 'loops' are involved (this number is called the *genus* of the surface): topologically, 0-loop (i.e. 'tree-level') world-sheets are spheres, 1-loop world-sheets are tori, etc. These closed Riemann surfaces are discussed in more detail in §2.3. The 0-loop contribution isn't very interesting (there is only one conformal equivalence class of spheres), so let us look at the 1-loop contribution. It will be of the form $\int \mathcal{Z}([\mathit{torus}])\, d[\mathit{torus}]$, where $\mathcal{Z}$ is a complex-valued function called the partition function, and $[\mathit{torus}]$ is a conformal equivalence class of tori. In the Segal formalism we recover this in the following way: the functor takes $[torus]$ to a linear function from $H^0 = \mathbb{C}$ to $H^0 = \mathbb{C}$. Any such linear function is simply a 1×1 matrix, i.e. a complex number, which we call $\mathcal{Z}([torus])$.

Now there is a nice parametrisation of conformal equivalence classes of tori, as we will see more explicitly in §2.3. Namely, a representative for each class can be chosen to be of the form $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$ where τ is in the upper half plane $\mathcal{H}$. Thus we can write $\mathcal{Z}$ as a function of a complex variable τ. However, different τ correspond to the same equivalence class of tori: the redundancy is exactly captured by the modular group $\mathrm{PSL}_2(\mathbb{Z})$. Namely, τ and $\frac{a\tau+b}{c\tau+d}$ are equivalent, whenever $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. Thus $\mathcal{Z}(\tau) = \mathcal{Z}(\frac{a\tau+b}{c\tau+d})$. In other words, the partition function is modular invariant![1]

There are two sectors in CFT, a holomorphic one and an antiholomorphic one, corresponding to the two directions ('left-' and 'right-moving') of motion on a string, or the two components of the group $\mathrm{Diff}(S^1)$ of diffeomorphisms of the circle. This means that many of the quantities (e.g. the partition function) factorise into parts depending holomorphically and anti-holomorphically on the modular parameters (e.g. τ in genus 1). In a *rational* CFT there are finitely many 'primary fields' $a \in \Phi$ — the precise meaning of this is not important here, but it says that the space of states for the theory decomposes into a finite sum[2] $H = \oplus_{a,b\in\Phi} M_{ab} H_a \otimes \overline{H}_b$, where M_{ab} are nonnegative integers which count the multiplicity of $H_a \otimes \overline{H}_b$ in H. The linear maps $\mathcal{T}(\Sigma) : H^m \to H^n$ in an RCFT will factorise similarly; this 'chiral factorisation' is captured by what Segal calls the 'modular functor' [52]. The partition function becomes

$$\mathcal{Z}(\tau) = \sum_{a,b\in\Phi} M_{ab}\, \chi_a(\tau)\, \chi_b(\tau)^* \tag{1.1.1}$$

for certain holomorphic functions χ_a. One of the primary fields (we'll denote it '0') corresponds to the vacuum Ω, and uniqueness of the vacuum means that $M_{00} = 1$.

H_0 is called a *chiral algebra*; in the language of §2.6, H_0 will be a vertex operator algebra (VOA). Φ parametrises the irreducible H_0-modules and the χ's are their characters; in an RCFT we require this number to be finite. For example, for the Moonshine VOA $V^\natural$ discussed in Part 2, Φ consists of only one element.

The higher-genus behaviour of an RCFT is determined from the lower-genus behaviour, by composition of 'arrows' (i.e. the gluing together of surfaces) in **C**. See Figure 3 of [30] to find how a genus-2 surface is built up from genus-0 ones. In fact, it's generally believed that an RCFT will be uniquely determined by: (i) the choice of chiral algebra; (ii) the partition function (which tells you the spectrum of the theory, i.e. how the two sectors link up); and (iii) the structure constants C_{ab}^c, which in the Segal formalism correspond to the surfaces called 'pairs-of-pants', equivalently disks with two interior disks removed. Our approach will be to start with a chiral algebra, and find all possible partition functions. We will thus ignore the important question of existence and uniqueness of the structure constants, though at least for our chiral algebras, it seems to be generally believed that the structure constants will be unique.

---

[1] In higher-dimensional string theories, a similar argument shows more generally that *automorphic forms* will appear naturally.

[2] It seems though that 'rational' logarithmic CFT is trying to teach us the lesson that this familiar requirement can and should be weakened. See Gaberdiel-Kausch (1999).

Perhaps all chiral algebras come from standard constructions (e.g. orbifolds and the Goddard-Kent-Olive (GKO) coset construction — see e.g. [18]) involving lattices and affine Kac-Moody algebras. For instance a $\mathbb{Z}_2$-orbifold of the VOA of the Leech lattice gives us the Moonshine module $V^\natural$, and the so-called minimal models come from GKO cosets involving $A_1^{(1)}$. This is in line with the spirit of Tannaka-Krein duality (and its generalisations by Deligne and Doplicher-Roberts), which roughly says that if a bunch of things *act like* they're the set of representations of a Lie group, then they *are* the set of representations of a Lie group.

In any case, one of the simplest, best understood, and important classes (called Wess-Zumino-Witten (WZW) models — see for instance [30,59] in this volume) of RCFTs correspond to affine Kac-Moody algebras at a positive integer level k. We will have much more to say later about these algebras, but for now let us remark that Φ here will be the (finite) set P_+^k of integrable level k highest weights λ. Their chiral algebras were constructed by Frenkel and Zhu. The following sections concern the attempt to classify the partition functions corresponding to Kac-Moody algebras — see especially §1.5. I will use this theme as an excuse to describe many other things, e.g. the A-D-E meta-pattern, Lie theory, Galois, fusion rings, ... I dedicate these notes to the profound friendship developing in recent years between mathematics and physics. As Victor Kac said in his 1996 Wigner medal acceptance speech, "Some of the best ideas come to my field from the physicists. And on top of this they award me a medal. One couldn't hope for a better deal."

1.2. LIE ALGEBRAS

Lie algebras (and their nonlinear partners *Lie groups*) appear in numerous places throughout math and mathematical physics. A nice introduction is [9]; Lie theory is presented with more of a physics flavour in [24], as well as [59].

An *algebra* is a vector space with a way to multiply vectors which is compatible with the vector space structure (i.e. the vector-valued product is *bilinear*: $(a\vec{u}+a'\vec{u}')\,(b\vec{v}+b'\vec{v}') = ab\,\vec{u}\vec{v} + ab'\,\vec{u}\vec{v}' + a'b\,\vec{u}'\vec{v} + a'b'\,\vec{u}'\vec{v}')$. For example, the complex numbers $\mathbb{C}$ can be thought of as a 2-dimensional algebra over $\mathbb{R}$ (a basis is 1 and $\mathrm{i} = \sqrt{-1}$; the *scalars* here are real numbers and the *vectors* are complex numbers). The quaternions are 4-dimensional over $\mathbb{R}$ and the octonions are 8-dimensional over $\mathbb{R}$. Incidentally, these are the only finite-dimensional algebras over $\mathbb{R}$ which obey the cancellation law: $\vec{u} \neq 0$ and $\vec{u}\vec{v} = 0$ implies $\vec{v} = 0$ (the reader should try to convince himself why the familiar vector product on $\mathbb{R}^3$ fails the cancellation law). This important little fact makes several unexpected appearances in math. For instance, it is trivially possible to 'comb the hair' on the circle S^1 without 'cheating' (i.e. needing a hair-part or exploiting a bald spot): just comb the hair clockwise for example. However it is not possible to comb the hair on the sphere S^2 (e.g. your own head) without cheating. The only other k-spheres S^k which can be combed (i.e. for which there exist k linearly independent continuous vector fields) are $k = 3$ and 7. This is intimately connected with the existence of $\mathbb{C}$, the quaternions, and octonions (namely,

S^1, S^3, S^7 can be thought of as the length 1 complex numbers, quaternions, and octonions, resp.).

In a *Lie* algebra $\mathfrak{g}$, the product is usually called a 'bracket' and is written $[xy]$. It is required to be 'anti-commutative' and 'anti-associative':

$$[xy] + [yx] = 0 \tag{1.2.1a}$$

$$[x[yz]] + [y[zx]] + [z[xy]] = 0 \tag{1.2.1b}$$

(like most other equalities in math, (1.2.1b) is usually called the *Jacobi identity*). Usually we will consider Lie algebras over $\mathbb{C}$, but sometimes over $\mathbb{R}$. Note that (1.2.1a) says $[xx] = 0$.

One important consequence of bilinearity is that it is enough to know the values of all the brackets $[x_i x_j]$ for $i < j$, for any basis $\{x_1, x_2, \ldots\}$ of $\mathfrak{g}$. (The reader should convince himself of this before proceeding.)

The simplest example of a Lie algebra is $\mathfrak{g} = \mathbb{C}$ (or $\mathfrak{g} = \mathbb{R}$), with the bracket $[xy]$ identically 0. In fact, this is the only 1-dimensional Lie algebra. It is a straightforward exercise for the reader to find all 2- and 3-dimensional Lie algebras (over $\mathbb{C}$) up to isomorphism (i.e. change of basis): there are precisely 2 and 6 of them, respectively (though one of the 6 depends on a complex parameter). Over $\mathbb{R}$, there are 2 and 9 (with 2 depending on real parameters). This exercise cannot be continued much further — e.g. not all 7-dimensional Lie algebras (over $\mathbb{C}$) are known. Nor is it obvious that this would be an interesting or valuable exercise. We should suspect that our definition of Lie algebra is probably a little too general for anything obeying it to be automatically an interesting structure. More often than not, a classification turns out to be a stale and useless list.

Two of the 3-dimensional Lie algebras are important in what follows. One of them is well-known to the reader: consider the vector-product (also called cross-product) in $\mathbb{C}^3$. Taking the standard basis $\{e_1, e_2, e_3\}$ of $\mathbb{C}^3$, the bracket can be defined by the relations

$$[e_1 e_2] = e_3 \ , \qquad [e_1 e_3] = -e_2 \ , \qquad [e_2 e_3] = e_1 \ . \tag{1.2.2a}$$

This Lie algebra, denoted A_1 or $\mathrm{sl}_2(\mathbb{C})$, can be called the 'mother of all (semi-simple) Lie algebras'. A more familiar realisation of A_1 uses a basis $\{e, f, h\}$ with relations

$$[ef] = h \ , \qquad [he] = 2e \ , \qquad [hf] = -2f \ . \tag{1.2.2b}$$

The reader can find the change-of-basis (valid over $\mathbb{C}$ but not $\mathbb{R}$) showing that (1.2.2) define isomorphic *complex* (but not *real*) Lie algebras.

Another important 3-dimensional Lie algebra is called the Heisenberg algebra[3] and is the algebra of the canonical commutation relations in quantum mechanics: choosing a basis x, p, h, it is defined by

$$[xp] = h \ , \qquad [xh] = [ph] = 0 \ . \tag{1.2.3}$$

---

[3] Actually, 'Heisenberg algebra' refers to a family of Lie algebras, with (1.2.3) being the one of lowest dimension.

From our definition, it is far from clear that Lie algebras, as a class, should be natural and worth studying. After all, there are infinitely many possible axiomatic systems: why should the one defining a Lie algebra be anything special *a priori*? Perhaps this could have been anticipated by the following line of reasoning.

Axiom. Groups are important and interesting.
Axiom. Manifolds are important and interesting.

Manifolds are structures where calculus is possible; locally, a manifold looks like a piece of $\mathbb{R}^n$ (or $\mathbb{C}^n$), but these pieces can be bent and stitched together to create more interesting shapes. For instance a circle is a 1-dimensional manifold, while Einstein claimed space-time is a curved 4-dimensional one.

Definition. A Lie group is a manifold with a compatible group structure.

This means that 'multiplication' and 'inverse' are differentiable maps. $\mathbb{R}$ is a Lie group, under addition: obviously, $\mu : \mathbb{R}^2 \to \mathbb{R}$ and $\iota : \mathbb{R} \to \mathbb{R}$ defined by $\mu(a,b) = a + b$ and $\iota(a) = -a$ are both differentiable. (Why isn't $\mathbb{R}$ a Lie group under multiplication?) A circle is also a Lie group: parametrise the points with the angle θ defined mod 2π (or mod 360 if you prefer); the 'product' of the point at angle θ_1 with the point at angle θ_2 will be the point at angle $\theta_1 + \theta_2$. Surprisingly, the only other k-sphere which is a Lie group is S^3 (the product can be defined using quaternions of unit length[4], or using the matrix group $\mathrm{SU}_2(\mathbb{C})$). Many but not all Lie groups can be expressed as matrix groups. Two other examples are GL_n (invertible $n \times n$ matrices) and SL_n (ones with determinant 1).

A consequence of the above axioms is then surely:

Corollary. Lie groups should be important and interesting.

Lie group structure theory can be thought of as a major generalisation of linear algebra. The basic constructions familiar to undergraduates have important analogues valid in many Lie groups. For instance, years ago we were taught to solve linear equations and invert matrices by using elementary row operations to reduce a matrix to row-echelon form. What this says is that any matrix $A \in \mathrm{GL}_n(\mathbb{C})$ can be factorised $A = BPN$, where N is uppertriangular with 1's on the diagonal, P is a permutation matrix, and B is an uppertriangular matrix. This is essentially what is called the Bruhat decomposition of the Lie group $\mathrm{GL}_n(\mathbb{C})$. More generally (where it applies to any 'reductive' Lie group G), P will be an element of the so-called 'Weyl group' of G (of which we'll have much more to say later), and B will be in a 'Borel subgroup'.

Lie groups appear throughout physics. E.g. the orthogonal group $\mathrm{SO}_3(\mathbb{R})$ is the configuration space of a rigid body centred at the origin, while $\mathrm{SU}_2(\mathbb{C})$ is the set of states of an electron at rest. The gauge group of the Standard Model of particle physics is $\mathrm{SU}_3(\mathbb{C}) \times \mathrm{SU}_2(\mathbb{C}) \times \mathrm{U}_1(\mathbb{C})$, while the Lorentz group of special relativity is $\mathrm{SO}_{3,1}(\mathbb{R})$.

There is an important relation between Lie groups and Lie algebras.

Fact. The tangent space of a Lie group is a Lie algebra. Any (finite-dimensional real or complex) Lie algebra is the tangent space to some Lie group.

---

[4] Similarly, the 7-sphere inherits from the octonions a *nonassociative* (hence nongroup) product, compatible with its manifold structure.

More precisely, the tangent space at 1 (i.e. the set of all infinitesimal generators of the Lie group) can be given a natural Lie algebra structure. A Lie algebra, being a linearised Lie group, is much simpler and easier to handle. The Lie algebra preserves the local properties of the Lie group, though it loses global topological properties (like boundedness). A Lie group has a single Lie algebra, but a Lie algebra can correspond to many different Lie groups. The Lie algebra corresponding to both $\mathbb{R}$ and S^1 is $\mathfrak{g} = \mathbb{R}$ with trivial bracket. The Lie algebra corresponding to both $S^3 = \mathrm{SU}_2(\mathbb{C})$ and $\mathrm{SO}_3(\mathbb{R})$ is the cross-product algebra on $\mathbb{R}^3$ (usually called $\mathrm{so}_3(\mathbb{R})$). Given the above fact, a safe guess would be:

Conjecture. Lie algebras are important and interesting.

From this line of reasoning, it should be expected that historically Lie groups arose first. Indeed that is the case: the Norwegian Sophus Lie introduced them in 1873 to try to develop a Galois theory for ordinary differential equations. As the reader may be aware, Galois theory is used for instance to show that not all 5th degree (or higher) polynomials can be explicitly 'solved' using radicals — we will meet Galois theory in §1.8. Lie wanted to study the explicit solvability (integrability) of differential equations, and this led him to develop what we now call Lie theory. The importance of Lie groups however have grown well beyond this initial motivation.

An important class of Lie algebras are the so-called finite-dimensional *simple* ones. Their definition and motivation will be studied in §2.7 below, but in a certain sense they serve as building blocks for all other finite-dimensional Lie algebras.

The classification of simple finite-dimensional Lie algebras over $\mathbb{C}$ is quite important and was accomplished at the turn of the century by Killing and Cartan. There are 4 infinite families A_ℓ ($\ell \geq 1$), B_ℓ ($\ell \geq 3$), C_ℓ ($\ell \geq 2$), and D_ℓ ($\ell \geq 4$), and 5 exceptionals E_6, E_7, E_8, F_4 and G_2. A_ℓ can be thought of[5] as $\mathrm{sl}_{\ell+1}(\mathbb{C})$, the $(\ell+1)\times(\ell+1)$ matrices with trace 0. The orthogonal algebras B_ℓ and D_ℓ can be identified with $\mathrm{so}_{2\ell+1}(\mathbb{C})$ and $\mathrm{so}_{2\ell}(\mathbb{C})$, resp., where $\mathrm{so}_n(\mathbb{C})$ is all $n \times n$ anti-symmetric matrices $A^t = -A$. The symplectic algebra C_ℓ is $\mathrm{sp}_{2\ell}(\mathbb{C})$, i.e. all $2\ell \times 2\ell$ matrices A obeying $A\Omega = -\Omega A^t$, where $\Omega = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$ and I_ℓ is the identity matrix. The exceptionals can be constructed e.g. using the octonions. In all these cases the bracket is given by the commutator

$$[AB] = [A, B] := AB - BA \tag{1.2.4}$$

(it is a good exercise for the reader to confirm that the commutator satisfies (1.2.1), and that e.g. $\mathrm{sl}_n(\mathbb{C})$ is indeed closed under it). To see that (1.2.2b) truly is $\mathrm{sl}_2(\mathbb{C})$, put

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.2.5}$$

Incidentally the names A, B, C, D have no significance: since the 4 series start at $\ell = 1, 2, 3, 4$, it seemed natural to call these A, B, C, D, resp. Unfortunately a bit of bad luck happened: B_2 and C_2 are isomorphic and so at random that algebra was placed in the orthogonal series; however affine Dynkin diagrams make it clear that it really is a symplectic algebra which accidentally looks orthogonal; hence in hindsight the names of the B- and C-series really should have been switched.

---

[5] Strictly speaking these are *representations* (see next section).

This classification changes if the *field* — the choice of scalars=numbers — is changed. By a field, we mean we can add, subtract, multiply and divide, such that all the usual properties like commutativity and distributivity are obeyed. Fields will make a few different appearances in these notes. $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Q}$ are fields, while $\mathbb{Z}$ is not (you can't always divide an integer by e.g. 3, and remain in $\mathbb{Z}$). The integers mod n, which we will write $\mathbb{Z}_n$, are a field iff n is prime (the reader can verify that in e.g. $\mathbb{Z}_4$, it is not possible to divide by the field element $[2] \in \mathbb{Z}_4$ even though $[2] \neq [0]$ there). $\mathbb{C}$ and $\mathbb{R}$ are examples of fields of characteristic 0 — this means that 0 is the only integer k with the property that $kx = 0$ for all x in the field. $\mathbb{Z}_p$ is the simplest example of a field with nonzero characteristic: in $\mathbb{Z}_p$, multiplying by the integer p has the same effect as multiplying by 0, and so we say $\mathbb{Z}_p$ has characteristic p. Strange fields have important applications in e.g. coding theory and, ironically, in number theory itself — see e.g. §1.8.

As always, $\mathbb{C}$ is better behaved than e.g. $\mathbb{R}$ because every polynomial can be factorised over $\mathbb{C}$ (we say $\mathbb{C}$ is *algebraically closed*) — this implies for example that every matrix has an eigenvector over $\mathbb{C}$ but not necessarily over $\mathbb{R}$. Over $\mathbb{R}$, the difference in the simple Lie algebra classification is that each symbol $X_\ell \in \{A_\ell, \ldots, G_2\}$ corresponds to a number of inequivalent algebras (over $\mathbb{C}$, each algebra has its own symbol). For example, 'A_1' corresponds to 3 different real simple Lie algebras, namely the matrix algebras $\mathrm{sl}_2(\mathbb{R})$, $\mathrm{sl}_2(\mathbb{C})$ (interpreted as a *real* vector space), and $\mathrm{so}_3(\mathbb{R})$. The simple Lie algebra classification has recently been done in any characteristic $p > 7$. It is surprising but very common that the smaller primes behave very poorly, and the classification for characteristic 2 is probably completely hopeless.

Associated with each simple algebra X_ℓ is a Weyl group, and a (Coxeter-)Dynkin diagram. The Weyl group is a finite reflection group, e.g. for A_ℓ it is the symmetric group $\mathfrak{S}_{\ell+1}$. See Figure 7 of [59] for the Weyl group of A_2. The Dynkin diagram of X_ℓ (see e.g. [24,36,38] or Figure 6 in [59]) is a graph with ℓ nodes, and with possibly some double and triple edges. It says how to construct X_ℓ abstractly using generators and relations — see §2.7. We will keep meeting both throughout these notes.

Another source of Lie algebras are the vector fields $\mathrm{Vect}(M)$ on a manifold M. A vector field v is a choice (in a smooth way) of a tangent vector $v(p) \in T_pM$ at each point of M. It can be thought of as a (1st order) differential operator, acting on functions $f : M \to \mathbb{R}$ (or $f : M \to \mathbb{C}$); at each point $p \in M$ take the directional derivative of f in the direction $v(p)$. For example the vector fields on the circle, $\mathrm{Vect}(S^1)$, can be thought of as anything of the form $g(\theta)\frac{d}{d\theta}$ where $g(\theta)$ can be any function with period 1. We can compose vector fields $u \circ v$, but this will result in a 2nd order differential operator: e.g.

$$(f(\theta)\frac{d}{d\theta}) \circ (g(\theta)\frac{d}{d\theta}) = f(\theta)\, g(\theta)\, \frac{d^2}{d\theta^2} + f(\theta)\, g'(\theta)\, \frac{d}{d\theta}\ .$$

Instead, the natural 'product' of vector fields is given by their commutator $[u, v] = u \circ v - v \circ u$, as it always results in a vector field: e.g.

$$[f(\theta)\frac{d}{d\theta}, g(\theta)\frac{d}{d\theta}] = (f(\theta)\, g'(\theta) - f'(\theta)\, g(\theta))\frac{d}{d\theta}$$

in $\mathrm{Vect}(S^1)$. $\mathrm{Vect}(M)$ with this bracket is an infinite-dimensional Lie algebra. In the case where M is a Lie group G, the Lie algebra of G can be interpreted as a certain finite-dimensional subalgebra of $\mathrm{Vect}(G)$ given by the 'left-invariant vector fields'.

Simple algebras need not be finite-dimensional. An example of an infinite-dimensional one is the *Witt algebra* $\mathcal{W}$, which can be defined (over $\mathbb{C}$) by the basis[6] L_n, $n \in \mathbb{Z}$, and the relations

$$[L_m L_n] = (m-n)L_{m+n} \ . \tag{1.2.6}$$

Using the realisation $L_n = -\mathrm{i}e^{-\mathrm{i}n\theta}\frac{d}{d\theta}$, the Witt algebra can also be interpreted as the polynomial subalgebra of the complexification $\mathbb{C} \otimes \mathrm{Vect}(S^1)$ — i.e. change the scalar field of $\mathrm{Vect}(S^1)$ from $\mathbb{R}$ to $\mathbb{C}$. Incidentally, infinite-dimensional Lie algebras need not have a corresponding Lie group: e.g. the real algebra $\mathrm{Vect}(S^1)$ is the Lie algebra of the Lie group $\mathrm{Diff}^+(S^1)$ of orientation-preserving diffeomorphisms $S^1 \to S^1$, but $\mathbb{C} \otimes \mathrm{Vect}(S^1)$ has no Lie group. $\mathrm{Diff}^+(S^1)$ plays a large role in CFT, by acting on the objects of Segal's category $\mathbf{C}$.

The Witt algebra appears naturally in CFT: e.g. using the realisation $L_n = -z^{n+1}\frac{d}{dz}$ it is the polynomial subalgebra of the Lie algebra $\mathrm{Vect}(\mathbb{C}/\{0\})$. Very carelessly, $\mathrm{Vect}(\mathbb{C}/\{0\})$ is often thought of as the infinitesimal conformal transformations on a suitable neighbourhood of 0 (yet clearly L_{-2}, L_{-3}, ... are singular at 0!). Indeed the CFT literature is very sloppy when discussing the conformal group in 2-dimensions. The unfortunate fact is that, contrary to claims, *there is no infinite-dimensional conformal group* for $\mathbb{C} \cong \mathbb{R}^2$. The best we can do is the 3-dimensional group $\mathrm{PSL}_2(\mathbb{C})$ of Möbius transformations $z \mapsto \frac{az+b}{cz+d}$, which are orientation-preserving conformal transformations for the Riemann sphere $\mathbb{C} \cup \{\infty\}$. There seem to be 2 ways out of this rather embarrassing predicament. One is to argue that we are really interested in 'infinitesimal conformal invariance' in some meromorphic sense, so the full Witt algebra can appear. The other way is to argue that it is the conformal group of 'Minkowski space' $\mathbb{R}^{1,1}$ (or better, its compactification $S^1 \times S^1$) rather than $\mathbb{R}^2 \cong \mathbb{C}$ (or its compactification S^2) which is relevant for CFT. That conformal group *is* infinite-dimensional; for $S^1 \times S^1$ it consists of 2 copies of $\mathrm{Diff}^+(S^1) \times \mathrm{Diff}^+(S^1)$. For a more careful treatment of this point, see [51].

For reasons we will discuss in §1.4, we are more interested in the *Virasoro algebra* $\mathcal{V}$ rather than the Witt algebra $\mathcal{W}$. This is a '1-dimensional central extension' of $\mathcal{W}$; as a vector space $\mathcal{V} = \mathcal{W} \oplus \mathbb{C}C$ with relations given by

$$[L_m L_n] = (m-n)L_{m+n} + \delta_{n,-m}\frac{m\,(m^2-1)}{12}\,C \tag{1.2.7a}$$

$$[L_m C] = 0 \ . \tag{1.2.7b}$$

'1-dimensional central extension' means $\mathcal{V}$ has one extra basis vector C, which lies in the *centre* of $\mathcal{V}$ (i.e. $[xC] = 0$ for all $x \in \mathcal{V}$), and sending $C \to 0$ recovers $\mathcal{W}$ (i.e. takes (1.2.7a) to (1.2.6)). A common mistake in the physics literature is to regard C as a number: it is in fact a vector, though in many (but not all) representations it is mapped to a scalar multiple of the identity.

---

[6] In infinite dimensions, to avoid convergence complications, only finite linear combinations of basis vectors are generally permitted. Infinite linear combinations would involve taking some 'completion'.

The reason for the strange-looking (1.2.7a) is that we have little choice: $\mathcal{V}$ is the unique nontrivial 1-dimensional central extension of $\mathcal{W}$. The factor $\frac{1}{12}$ there is conventional but standard, and has to do with 'zeta-function regularisation' in string theory — i.e. the divergent sum $\sum_{n=1}^{\infty} n$ is 'reinterpreted' as $\zeta(-1) = \frac{-1}{12}$, where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ is the Riemann zeta function. Incidentally $\zeta(s)$ can be written as the product $\prod_p (1-p^{-s})^{-1}$ over all primes $p = 2, 3, 5, \ldots$ (try to see why); hence $\zeta(s)$ has a lot to do with primes, in particular their distribution. In fact the most famous unsolved problem in math today is the Riemann conjecture, which says that $\zeta(s) \neq 0$ whenever $\mathrm{Re}(s) \neq \frac{1}{2}$. One researcher recently described this conjecture as saying that the primes have music in them.

In CFT, L_0 is the energy operator. For example the partition function is given by $\mathcal{Z}(\tau) = \mathrm{Tr}_H(q^{L_0-c/24} q^{*\bar{L}_0-c/24})$ and the (normalised) character χ_a equals $\mathrm{Tr}_{H_a}(q^{L_0-c/24})$ for $q = e^{2\pi \mathrm{i}\tau}$. cI is the scalar multiple of the identity to which C gets sent; it has a physical interpretation [18] involving Casimir (vacuum) energy, which depends on space-time topology, and the strange shift by $c/24$ is due to an implicit mapping from the cylinder to the plane.

1.3. REPRESENTATIONS OF FINITE-DIMENSIONAL SIMPLE LIE ALGEBRAS

The representation theory of the simple Lie algebras[7] can perhaps be regarded as an enormous generalisation of trigonometry. For instance the facts that $\frac{\sin(nx)}{\sin(x)}$ can be written as a polynomial in $\cos(x)$ for any $n \in \mathbb{Z}$, and that

$$\frac{\sin(mx)\,\sin(nx)}{\sin(x)} = \sin((m+n)x) + \sin((m+n-2)x) + \cdots + \sin((m-n)x)$$

for any $m, n \in \mathbb{Z}_>$, are both easy special cases of the theory.

The classic example of an algebraic structure are the numbers, and they prejudice us into thinking that commutativity and associativity are the ideal. We have learned over the past couple of centuries that commutativity can often be dropped without losing depth and usefulness, but most interesting structures seem to obey some form of associativity. Moreover, true associativity (as opposed to e.g. anti-associativity) really simplifies the arithmetic. Given the happy 'accident' that the commutator $[x, y] := xy - yx$ in any associative algebra obeys anti-associativity, it would seem to be both tempting and natural to study the ways (if any) in which associative algebras $\mathfrak{A}$ can 'model' or *represent* a given Lie algebra. Precisely, we are looking for a map $\rho : \mathfrak{g} \to \mathfrak{A}$ which preserves the linear structure (i.e. ρ is a linear function), and which sends the bracket $[xy]$ in $\mathfrak{g}$ to the commutator $[\rho(x), \rho(y)]$ in $\mathfrak{A}$.

In practice groups (resp., algebras) often appear as symmetries (resp., infinitesimal generators of symmetries). These symmetries often act linearly. In other words, in practise the preferred associative algebras will usually be matrix algebras, and this is the usual form for a representation and the only kind we will consider. The *dimension* of these representations is the size of the matrices.

---

[7] See e.g. [25] for more details. Historically, representations of Lie algebras were considered even before representations of finite groups.

Finding all possible representations, even for the simple Lie algebras, is probably hopeless. However, it is possible to find all *finite-dimensional* representations of the simple Lie algebras, and the answer is easy to describe. Given a simple Lie algebra X_ℓ, there is a representation L_λ for each ℓ-tuple $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of nonnegative integers. λ is called a *highest-weight*. Moreover, we can take direct sums $\oplus_i L_{\lambda^{(i)}}$ of finitely many of these representations. The matrices in such a direct sum will be in block form. It turns out that, up to change-of-basis, this exhausts all finite-dimensional representations of X_ℓ.

It is common to replace 'representation ρ of $\mathfrak{g}$' with the equivalent notion of '$\mathfrak{g}$-module M' — i.e. we think of the matrices $\rho(x)$ as linear maps $M \to M$. A $\mathfrak{g}$-*module* is a vector space on which $\mathfrak{g}$ acts (on the left). Instead of considering the matrix $\rho(x)$, we consider 'products' xv (think of this as the matrix $\rho(x)$ times the column vector v) for $v \in M$. This product must be bilinear, and must obey $[xy]v = x(yv) - y(xv)$.

To get an idea of what L_λ looks like, consider A_1. Recall its generators e, f, h and relations (1.2.2b). Choose any $\lambda \in \mathbb{C}$. Define $x_0 \neq 0$ to formally obey $hx_0 = \lambda x_0$ and $ex_0 = 0$. Define inductively $x_{i+1} := fx_i$ for $i = 0, 1, \ldots$. Define M_λ to be the span of all x_i — we will see shortly that they are linearly independent (so M_λ is infinite-dimensional). M_λ is a module of A_1: the calculations $hx_{i+1} = hfx_i = ([hf] + fh)x_i = (-2f + fh)x_i$ and $ex_{i+1} = efx_i = ([ef] + fe)x_i = (h + fe)x_i$ show inductively that $hx_m = (\lambda - 2m)x_m$ and $ex_m = (\lambda - m + 1)m\, x_{m-1}$. From these the reader can show that the x_i are linearly independent. M_λ is called a *Verma module*; λ is called its highest-weight, and x_0 is called a highest-weight vector.

Now specialise to $\lambda = n \in \mathbb{Z}_\geq := \{0, 1, 2, \ldots\}$. Note that $ex_{n+1} = 0$ and $hx_{n+1} = (-n-2)x_{n+1}$. This means that, for these n, M_n contains a *submodule* with highest-weight vector x_{n+1}, isomorphic to M_{-n-2}. x_{n+1} is called a *null vector*. In other words, we could set $x_{n+1} := 0$ and still have an A_1-module. We would then get a *finite-dimensional* module which we'll call $L_n := M_n/M_{-n-2}$ (not to be confused with the Virasoro generator in (1.2.7)). Its basis is $\{x_0, x_1, \ldots, x_n\}$ and so it has dimension $n + 1$.

For example, take $n = 1$. Note that what we get in terms of the basis $\{x_0, x_1\}$ is the familiar representation $\mathrm{sl}_2(\mathbb{C})$ given in (1.2.5).

The situation for the other simple Lie algebras X_ℓ is similar.

It turns out to be hard to compare representations: ρ and ρ' could be equivalent (i.e. differ merely by a change-of-basis) but look very different. Or if we are given a representation, we may want to decompose it into the direct sum of some $L_{\lambda^{(i)}}$. When working with representations, it is often very useful to avoid much of the extraneous basis-dependent detail present in the function ρ. Finite group theory suggests how to do this: we should use *characters*. The character of an A_1-module M is given by Weyl: write M as a direct sum of eigenspaces $M(m)$ of h; then define

$$\mathrm{ch}_M(z) := \sum_m \dim\, M(m)\; e^{mz}\ , \tag{1.3.1}$$

for any $z \in \mathbb{C}$. The m are called *weights* and the $M(m)$ *weight-spaces*. For example, for

L_n the weights are $m = n, n-2, \ldots, -n$, the weight-spaces $L_n(m)$ are $\mathbb{C}\, x_{(n-m)/2}$, and

$$\mathrm{ch}_n(z) = \sum_{i=0}^{n} e^{(n-2i)z} = \frac{\sin((n+1)\,z)}{\sin(z)}\ . \tag{1.3.2}$$

Analogous formulas apply to any algebra X_ℓ: the character will then be a function of an ℓ-dimensional subspace $\mathfrak{h}$ called the *Cartan subalgebra*, spanned by all the h_i (see §2.7), so can be thought of as a complex-valued function of ℓ complex variables. The weights m will lie in the dual space to $\mathfrak{h}$ — i.e. are linear maps $\mathfrak{h} \to \mathbb{C}$ — so will have ℓ components. See for instance Figure 8 in [59]. Incidentally, ℓ is called the *rank* of X_ℓ.

Weyl's definition works: two representations are equivalent iff their characters are identical, and $M = \oplus_i L_{\lambda^{(i)}}$ iff $\mathrm{ch}_M(z) = \sum_i \mathrm{ch}_{\lambda^{(i)}}(z)$. It also is enormously simpler: e.g. the smallest nontrivial representation of E_8 is a map from $\mathbb{C}^{248}$ to the space of 248×248 matrices, while its character is a function $\mathbb{C}^8 \to \mathbb{C}$. But why is Weyl's definition natural? How did he come up with it?

To answer that question, we must remind ourselves of the characters of finite groups[8]. A representation of a finite group G is a structure-preserving map ρ (i.e. a group homomorphism) from G to matrices. The group's product becomes matrix product. In these notes we will be exclusively interested in group representations over $\mathbb{C}$. Two representations ρ, ρ' are called *equivalent* if there exists a matrix (change-of-basis) U such that $\rho'(g) = U\rho(g)U^{-1}$ for all g. The character ch_ρ is the map $G \to \mathbb{C}$ given by the trace: $\mathrm{ch}_\rho(g) = \mathrm{tr}(\rho(g))$. We see that equivalent representations will have the same character, because of the fundamental identity $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. This identity also tells us that the character is a 'class function', i.e. $\mathrm{ch}_\rho(hgh^{-1}) = \mathrm{tr}(\rho(h)\,\rho(g)\,\rho(h)^{-1}) = \mathrm{ch}_\rho(g)$ so ch_ρ is constant on each 'conjugacy class'. Group characters are also enormously simpler than representations: e.g. the smallest nontrivial representation of the Monster $\mathbb{M}$ (see Part 2) consists of almost 10^{54} matrices, each of size 196883×196883, while its character consists of 194 complex numbers. Incidentally, finite group representations behave analogously to the representations of X_ℓ: the role of the modules L_λ is played by the irreducible representations ρ_i, and any finite-dimensional representation of G can be decomposed uniquely into a direct sum of various ρ_i. The difference is that there are only finitely many ρ_i — their number equals the number of conjugacy classes of G.

We can use this group intuition here. In particular, given any Lie algebra X_ℓ and representation ρ, we can think of the map $e^x \mapsto e^{\rho(x)}$ as a representation of a Lie group $G(X_\ell)$ corresponding to X_ℓ (the exponential e^A of a matrix is defined by the usual power series; it will always converge). The trace of the matrix $e^{\rho(x)}$ will be the *group* character value at $e^x \in G(X_\ell)$, so we'll define it to be the *algebra* character value at $x \in X_\ell$. Again, it suffices to consider only representatives of each conjugacy class of $G(X_\ell)$, because the character will be a class function. Now, almost every matrix is diagonalisable (since almost any $n \times n$ matrix has n distinct eigenvalues), and so it would seem we aren't losing much by restricting $x \in X_\ell$ to *diagonalisable* matrices. Hence we may take our conjugacy class representatives to be *diagonal* matrices $x \in X_\ell$, i.e. (for $X_\ell = A_1$) to $x = zh$ for $z \in \mathbb{C}$ (h is diagonal in the x_i basis of L_λ). So the algebra character can be chosen to be a function of z. Finally, the trace of $e^{\rho(x)} = e^{z\rho(h)}$ will be given by (1.3.1). This completes the motivation for Weyl's character formula.

---

[8] Surprisingly, what we now call the characters of group representations were invented almost a decade before group representations were.

There is one other important observation we can make. Different diagonal matrices can belong to the same conjugacy class. For instance,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix} ,$$

so e^{zh} and e^{-zh} lie in the same $G(A_1) = \mathrm{SL}_2(\mathbb{C})$ conjugacy class. Hence $\mathrm{ch}_M(z) = \mathrm{ch}_M(-z)$. This symmetry $z \mapsto -z$ belongs to the *Weyl group* for A_1. Each X_ℓ has similar symmetries, and the Weyl group plays an important role in the whole theory, sort of analogous to the modular group for modular functions we'll discuss in §2.3.

Weyl found a generalisation of the right-side of (1.3.2), valid for all X_ℓ. The character of L_λ can be written as a fraction (2.8.1): the numerator will be a alternating sum over the Weyl group, and the denominator will be a product over 'positive roots'. This formula and its generalisations have profound consequences, as we'll see in §2.8.

Incidentally, the trigonometric identities given at the beginning of this section are the tensor product formula of representations (interpreted as the product and sum of characters), and the fact that an arbitrary character can be written as a polynomial in the fundamental characters, both specialised to A_1 (see (1.3.2) for the A_1 characters).

1.4. AFFINE ALGEBRAS AND THE KAC-PETERSON MATRICES

The theory of nontwisted affine Kac-Moody algebras (usually called *affine algebras* or *current algebras*) is extremely analogous to that of the finite-dimensional simple Lie algebras. Nothing infinite-dimensional tries harder to be finite-dimensional than affine algebras. Standard references for the following material are [38,41,24].

Let X_ℓ be any simple finite-dimensional Lie algebra. The affine algebra $X_\ell^{(1)}$ is essentially the *loop algebra* $\mathcal{L}(X_\ell)$, defined to be all possible 'Laurent polynomials' $\sum_{n\in\mathbb{Z}} a_n t^n$ where each $a_n \in X_\ell$ and all but finitely many $a_n = 0$. t here is an indeterminant. The bracket in $\mathcal{L}(X_\ell)$ is the obvious one: e.g. $[at^n, bt^m] = [ab]t^{n+m}$. Geometrically, $\mathcal{L}(X_\ell)$ is the Lie algebra of polynomial maps $S^1 \to X_\ell$ — hence the name (for that realisation, think of $t = e^{2\pi \mathrm{i}\theta}$). Hence there are many generalisations of the loop algebra (e.g. any manifold in place of S^1 will do), closely related ones called *toroidal algebras* being the Lie algebra of maps $S^1 \times \cdots \times S^1 \to X_\ell$. But the loop algebra is simplest and best understood, and the only one we'll consider. Note that $\mathcal{L}(X_\ell)$ is infinite-dimensional. Its Lie groups are the *loop groups*, consisting of all loops $S^1 \to G(X_\ell)$ in a Lie group for X_ℓ.

We saw S^1 before, in the discussion of the Witt algebra. Thus the Virasoro and affine algebras should be related. In fact, the Virasoro algebra acts on the affine algebras as 'derivations', and this connection plays an important technical role in the theory.

$X_\ell^{(1)}$ is in the same relation to the loop algebra, that the Virasoro $\mathcal{V}$ is to the Witt $\mathcal{W}$. Namely, it is its (unique nontrivial 1-dimensional) central extension — see e.g. (7.7.1) of [38] for the analogue of (1.2.7a) here. In addition, for more technical reasons, a further (noncentral) 1-dimensional extension is usually made: the derivation $t\frac{d}{dt}$ is included (see footnote 33). $X_\ell^{(1)}$ is the simplest of the infinite-dimensional Kac-Moody algebras. The superscript '(1)' denotes the fact that the loop algebra was twisted by an order-1 automorphism — i.e. that it is untwisted. It is called 'affine' because of its Weyl group, as we shall see.

Central extensions are a common theme in today's infinite-dimensional Lie theory[9]. Their *raison d'être* is always the same: a richer supply of representations. For example, $\mathcal{W}$ has several representations, but no nontrivial one is an 'irreducible unitary positive-energy representation' — the kind of greatest interest in math phys. On the other hand, its central extension $\mathcal{V}$ has a rich supply of those representations (e.g. there's one for each choice of $c > 1, h > 0$, namely the Verma module $V_{c,h}$ corresponding to $L_0 x_0 = hx_0, Cx_0 = cx_0$). At the level of groups, central extensions allow *projective* representations (i.e. representations up to a scalar factor) to become true representations. Projective representations (hence central extensions) appear naturally in QFT because a quantum state vector $|v\rangle$ is physically indistinguishable from any nonzero scalar multiple $\alpha|v\rangle$.

All of the quantities associated to X_ℓ have an analogue here: Dynkin diagram, Weyl group, weights,... For instance, the affine Dynkin diagram is obtained from the Dynkin diagram for X_ℓ by adding one node. See for example Figure 9 of [59]. The extra node is always labelled by a '0'. The Cartan subalgebra $\mathfrak{h}$ here will be $(\ell+2)$-dimensional. Many of these details will be discussed in more detail in §2.7 below.

The construction of $X_\ell^{(1)}$ is so trivial that it seems surprising anything interesting and new can happen here. But a certain 'miracle' happens...

No interesting representation of $X_\ell^{(1)}$ is finite-dimensional. The analogue for $X_\ell^{(1)}$ of the finite-dimensional representations of X_ℓ are called the *integrable highest-weight representations*, and will be denoted L_λ. The *highest-weight* λ here will be an $(\ell+1)$-tuple $(\lambda_0, \lambda_1, \ldots, \lambda_\ell)$, $\lambda_i \in \mathbb{Z}_\geq$ (strictly speaking, it will be an $(\ell+2)$-tuple, but the extra component is not important and is usually ignored). As for X_ℓ, the highest-weights can be thought of as the assignment of a nonnegative integer to each node of the Dynkin diagram. The construction of L_λ is as in the finite-dimensional case. They are called *integrable* because they are precisely those highest-weight representations which can be 'integrated' to a projective representation of the corresponding loop group, and hence a representation of a central extension of the loop group.

We define the character χ_λ as in (1.3.1), though now the weights m will be $(\ell+2)$-tuples, and there will be infinitely many of them. χ_λ will be a complex-valued function of $\ell+2$ complex variables $(\vec{z}, \tau, u)$ (see (1.4.1a) below). It can be written as an alternating sum over the Weyl group W, over a 'nice' denominator. The difference here is that W is now infinite.

Perhaps most of the interest in affine algebras can be traced to the 'miracle' that their Weyl groups are a semidirect product $Q^\vee \rtimes \overline{W}$ of translations in a lattice $\mathbb{Q}^\vee$ (the ℓ-dimensional 'co-root lattice' of X_ℓ — see §1.6) with the (finite) Weyl group $\overline{W}$ of X_ℓ.

---

[9] Incidentally the *finite-dimensional* simple Lie algebras do not have nontrivial central extensions.

See Figure 10 of [59] for the Weyl group of $A_2^{(1)}$. 'Semidirect product'[10] means that any element of W can be written uniquely as (t, w) for some translation t and some $w \in \overline{W}$, and $(t, w) \circ (t', w') = (\text{stuff}, w \circ w')$.

One thing this implies is that χ_λ will be of the form 'theta function'/denominator. Theta functions are classically-studied modular forms (we will discuss these terms in §2.3), and thus the modular group $\mathrm{SL}_2(\mathbb{Z})$ will make an appearance! To make this more precise, consider the highest-weight $\lambda = (\lambda_0, \lambda_1)$ of $A_1^{(1)}$, and write $k = \lambda_0 + \lambda_1$. Then

$$\chi_\lambda = \frac{\Theta_{\lambda_1+1}^{(k+2)} - \Theta_{-\lambda_1-1}^{(k+2)}}{\Theta_1^{(2)} - \Theta_{-1}^{(2)}} \tag{1.4.1a}$$

where these functions all depend on 3 complex variables z, τ, u, and

$$\Theta_m^{(n)}(z, \tau, u) := e^{-2\pi i n u} \sum_{\ell \in \mathbb{Z} + \frac{m}{2n}} \exp[\pi i n \tau \ell^2 - 2\sqrt{2}\pi i n \ell z] \ . \tag{1.4.1b}$$

In (1.4.1a) we can see the alternating sum over the Weyl group of A_1 in the numerator (and denominator, since we've used the $A_1^{(1)}$ denominator identity in writing (1.4.1a)). For general $X_\ell^{(1)}$, the denominator will always be independent of λ, and the theta function (1.4.1b) will become a multidimensional one involving a sum over $Q^\vee$ shifted by some weight and appropriately rescaled. The (co-)root lattice of A_1 is $\sqrt{2}\mathbb{Z}$. The key variable in (1.4.1a) is the modular one τ, which will lie in the upper half complex plane $\mathcal{H}$ (in order to have convergence). In the applications to CFT, the other variables are often set to 0.

The number k introduced in (1.4.1a) plays an important role in the general theory. In the representation L_λ, the central term C will get sent to some multiple of the identity — the multiplier is labelled k and is called the *level* of the representation. For any $X_\ell^{(1)}$ there is a simple formula expressing the level k in terms of the highest-weight λ; e.g. for $A_\ell^{(1)}$ and $C_\ell^{(1)}$ it is given by $k = \lambda_0 + \lambda_1 + \cdots + \lambda_\ell$. Write P_+^k for the (finite) set of level k highest-weights (so the size of P_+^k for $A_\ell^{(1)}$ is $\binom{k+\ell}{\ell}$). An important weight in P_+^k is $(k, 0, \ldots, 0)$. We will denote this '0'. In RCFT it corresponds to the vacuum.

The modular group $\mathrm{SL}_2(\mathbb{Z})$ acts on the Cartan subalgebra $\mathfrak{h}$ of $X_\ell^{(1)}$ in the following way:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (\vec{z}, \tau, u) = \left(\frac{\vec{z}}{c\tau + d}, \frac{a\tau + b}{c\tau + d}, u - \frac{c\,\vec{z} \cdot \vec{z}}{2(c\tau + d)}\right)$$

Under this action, the characters χ_λ also transform nicely: in particular we find for any level k weight λ

$$\chi_\lambda\left(\frac{\vec{z}}{\tau}, \frac{-1}{\tau}, u - \frac{\vec{z} \cdot \vec{z}}{2\tau}\right) = \sum_{\mu \in P_+^k} S_{\lambda\mu}\, \chi_\mu(\vec{z}, \tau, u) \tag{1.4.2a}$$

$$\chi_\lambda(\vec{z}, \tau + 1, u) = \sum_{\mu \in P_+^k} T_{\lambda\mu}\, \chi_\mu(\vec{z}, \tau, u) \tag{1.4.2b}$$

---

[10] This is also discussed briefly in section 2.2.

where S and T are complex matrices called the *Kac-Peterson matrices.* S will always be symmetric and unitary, and has many remarkable properties as we shall see. Its entries are related to Lie group characters at elements of finite order (see (1.4.5) below). T is diagonal and unitary; its entries are related to the eigenvalues of the quadratic Casimir.

For example, consider $A_1^{(1)}$ at level k. Then S and T will be $(k+1)\times(k+1)$ matrices given by

$$S_{\lambda\mu}=\sqrt{\frac{2}{k+2}}\,\sin(\pi\,\frac{(\lambda_1+1)(\mu_1+1)}{k+2})\ ,\qquad T_{\lambda\mu}=\exp[\pi\mathrm{i}\,\frac{(\lambda_1+1)^2}{2(k+2)}-\frac{\pi\mathrm{i}}{4}]\,\delta_{\lambda,\mu}\ . \tag{1.4.3}$$

One important place S appears is the famous *Verlinde formula*

$$N_{\lambda\mu}^{\nu}=\sum_{\kappa\in P_+^k}\frac{S_{\lambda\kappa}\,S_{\mu\kappa}\,S_{\nu\kappa}^*}{S_{0,\kappa}} \tag{1.4.4}$$

for the fusion coefficients $N_{\lambda\mu}^{\nu}$ of the corresponding RCFT. We will investigate some consequences of this formula in a later section. The fusion coefficients for the affine algebras are well-understood; see e.g. Section 4 of [59] for their interpretation (usually called the Kac-Walton formula) as 'folded tensor product coefficients'.

We will see in §1.7 that symmetries of the extended Dynkin diagram have consequences for S and T (simple-currents, charge-conjugation). There is a 'Galois action' on S which we will discuss in §1.8. There is a strange property of S and T called *rank-level duality* (see e.g. [45]): the matrices for $A_\ell^{(1)}$ at level k are closely related to those of $A_{k-1}^{(1)}$ at level $\ell+1$, and similar statements hold for $B_\ell^{(1)}$, $C_\ell^{(1)}$ and $D_\ell^{(1)}$. Another reason S is mathematically interesting is the formula

$$\frac{S_{\lambda\mu}}{S_{0\mu}}=\mathrm{ch}_{\bar{\lambda}}(-2\pi\mathrm{i}\,\frac{\overline{\mu+\rho}}{k+h^\vee})\ . \tag{1.4.5}$$

The right-side is a character of X_ℓ, and $\bar{\lambda}=(\lambda_1,\ldots,\lambda_\ell)$ means ignore the extended node. ρ is the 'Weyl vector' $(1,1,\ldots,1)$ and $h^\vee$ is called the *dual Coxeter number* and is the level of ρ. For $A_\ell^{(1)}$, $h^\vee=\ell+1$. Of course the right-side can also be regarded as a character for a Lie group associated to X_ℓ, in which case the argument would have to be exponentiated and would correspond to an element of finite order in the group. These numbers (1.4.5) have been studied by many people (most extensively by Pianzola) and have some nice properties. For instance Moody-Patera (1984) have argued that exploiting them leads to some quick algorithms for computing e.g. tensor product coefficients. Kac [37] found a curious application for them: a Lie theoretic proof of 'quadratic reciprocity'!

Quadratic reciprocity is one of the gems of classical number theory. It tells us that the equations

$$x^2\equiv a\ (\mathrm{mod}\ b)$$
$$y^2\equiv b\ (\mathrm{mod}\ a)$$

are related; more precisely, for fixed a and b (for simplicity take them to both be primes $\neq 2$) the questions of whether there is a solution x to the first equation and a solution y to the second, are related. They will both have the same yes or no answer, unless $a \equiv b \equiv 3 \pmod 4$, in which case they will have opposite answers. E.g. take $a = 23$ and $b = 3$, then we know the first equation does not have a solution (since $a \equiv 2 \pmod 3$ and $x^2 \equiv 2$ doesn't have a solution mod 3), and hence the second equation must have a solution (indeed, $y = 7$ works). There are now many proofs for quadratic reciprocity, and Kac used Lie characters at elements of finite order to find another one.

What is interesting here is that Kac's proof uses only certain special weights for A_ℓ. The natural question is: is it possible to find any generalisations of quadratic reciprocity using other weights and algebras? Many generalisations of quadratic reciprocity are known; will generalising Kac's argument recover them, or will they perhaps yield new reciprocity laws? It seems no one knows.

The relation (1.4.5) is important because it connects finite-dimensional Lie data with infinite-dimensional Lie data. The 'conceptual arrow' can be exploited both ways: in the generalisations of the arguments of §1.9 to other algebras, (1.4.5) allows us to use our extensive knowledge of finite-dimensional algebras to squeeze out some information in the affine setting; but also it is possible to use the richer symmetries of the affine data to see 'hidden' symmetries in finite-dimensional data. For example it can be used (Gannon-Walton 1995) to find a sort of Galois symmetry of dominant weight multiplicities in X_ℓ, which would be difficult or impossible to anticipate without (1.4.5).

1.5. The classification of physical invariants

We are interested in the following classification problem. Choose any affine algebra $X_\ell^{(1)}$ and level $k \in \mathbb{Z}_\geq$. Find all matrices $M = (M_{\lambda\mu})_{\lambda,\mu\in P_+^k}$ such that

(P1) $MS = SM$ and $MT = TM$, where S, T are the Kac-Peterson matrices (1.4.2);
(P2) each entry $M_{\lambda\mu} \in \mathbb{Z}_\geq$;
(P3) $M_{00} = 1$.

Any such M, or equivalently the corresponding partition function $\mathcal{Z} = \sum_{\lambda,\mu} M_{\lambda\mu}\chi_\lambda\chi_\mu^*$, is called a *physical invariant.*

The first and most important classification of physical invariants was the Cappelli-Itzykson-Zuber A-D-E classification for $A_1^{(1)}$ at all levels k [8]. We will give their result shortly. This implies for instance the minimal model RCFT classification, as well as the $N = 1$ super(symmetric)conformal minimal models. The other classifications of comparable magnitude are $A_2^{(1)}$ for all k; $A_\ell^{(1)}$, $B_\ell^{(1)}$ and $D_\ell^{(1)}$ for all $k \leq 3$; $(A_1 \oplus A_1)^{(1)}$ for all levels (k_1, k_2); and $(u(1) \oplus \cdots \oplus u(1))^{(1)}$ for all (matrix-valued) levels k. See e.g. [29] for references. The most difficult of these classifications is for $A_2^{(1)}$, done by Gannon (1994).

In other words, very little in this direction has been accomplished in the 15 or so years this problem has existed. But this is not really a good measure of progress. The effort instead has been directed primarily towards the full classification; most of these partial results are merely easy spin-offs from that more serious and ambitious assault.

The proof in [8] was very complicated and followed the following lines. First, an explicit basis was found for the vector space (called the 'commutant') of all matrices obeying (P1). Then (P2) and (P3) were imposed. Unfortunately their proof was long and formidable. Others tried to apply their approach to $A_2^{(1)}$, but without success. The eventual proof for $A_2^{(1)}$ was completely independent of the [8] argument, and exploited more of the structure implicit in the problem. As the $A_2^{(1)}$ argument became more refined, it became the model for the general assault. In §1.9 we sketch this new approach.

From this more general perspective, of these completed classifications only the level 2 $B_\ell^{(1)}$ and $D_\ell^{(1)}$ ones will have any lasting value (the orthogonal algebras at level 2 behave very peculiarly, possess large numbers of exceptional physical invariants, and must be treated separately). The others behave more generically and will fall out as special cases once the more general classifications are concluded. Other classifications which should be straightforward with our present understanding are $C_2^{(1)}$ at all k; $G_2^{(1)}$ at all k; and $B_\ell^{(1)}$ and $D_\ell^{(1)}$ at $k=4$. The $C_2^{(1)}$ should be easiest and would imply the $C_\ell^{(1)}$ level 2, as well as the $B_\ell^{(1)}$ and $D_\ell^{(1)}$ level 5, classifications. A very safe conjecture is that the only *exceptional* physical invariants (we define this term in §1.7) for $C_2^{(1)}$ occur at $k=3,7,8,12$ — this is known to be true for all $k \leq 500$. $G_2^{(1)}$ would be more difficult but also much more valuable; its only known exceptionals occur at $k=3,4$, and these are the only exceptionals for $k \leq 500$, and a very safe conjecture is that there are no other $G_2^{(1)}$ exceptionals. $B_\ell^{(1)}$ and $D_\ell^{(1)}$ at level 4 will also be more difficult, but also would be valuable; less is understood about its physical invariants and there is a good chance new exceptionals exist there.

The most surprising thing about the known physical invariant classifications is that there so few surprises: almost every physical invariant is 'generic'. We will see that the symmetries of the extended Dynkin diagram give rise to general families of physical invariants. We will call any physical invariants which do not arise in these generic ways (i.e. using what are called simple-currents or conjugations), *exceptional*. Many exceptionals have been found, and now we are almost at the point where we can safely conjecture the complete list of physical invariants for $X_\ell^{(1)}$ at any k, for X_ℓ a simple algebra.

Unfortunately the classification for semi-simple algebras $X_{\ell_1} \oplus \cdots \oplus X_{\ell_s}$ does not reduce to the one for simple ones. In fact, *any explicit classification of the physical invariants for $X^{(1)}$, for all semi-simple X, would easily be one of the greatest accomplishments in the history of math,* for it would include as a small part such monumental things as an explicit classification of all positive-definite integral lattices. Thus we unfortunately cannot expect an *explicit* classification for the *semi*-simple algebras.

To make this discussion more concrete and explicit, consider $A_1^{(1)}$. For convenience drop λ_0, so $P_+^k = \{0,1,\ldots,k\}$. Write J for the permutation (called a simple-current) $Ja := k-a$. Then the complete list of physical invariants for $A_1^{(1)}$ is

$$\mathcal{A}_{k+1} = \sum_{a=0}^{k} |\chi_a|^2 \;, \qquad\qquad \text{for all } k \geq 1$$

$$\mathcal{D}_{\frac{k}{2}+2} = \sum_{a=0}^{k} \chi_a \, \chi_{J^a a}^* \;, \qquad\qquad \text{whenever } \frac{k}{2} \text{ is odd}$$

$$\begin{aligned}
\mathcal{D}_{\frac{k}{2}+2} &= |\chi_0+\chi_{J0}|^2+|\chi_2+\chi_{J2}|^2+\cdots+2|\chi_{\frac{k}{2}}|^2\ , &&\text{whenever } \frac{k}{2} \text{ is even}\\
\mathcal{E}_6 &= |\chi_0+\chi_6|^2+|\chi_3+\chi_7|^2+|\chi_4+\chi_{10}|^2\ , &&\text{for } k=10\\
\mathcal{E}_7 &= |\chi_0+\chi_{16}|^2+|\chi_4+\chi_{12}|^2+|\chi_6+\chi_{10}|^2\\
&\quad +\chi_8\,(\chi_2+\chi_{14})^*+(\chi_2+\chi_{14})\,\chi_8^*+|\chi_8|^2\ , &&\text{for } k=16\\
\mathcal{E}_8 &= |\chi_0+\chi_{10}+\chi_{18}+\chi_{28}|^2+|\chi_6+\chi_{12}+\chi_{16}+\chi_{22}|^2\ , &&\text{for } k=28\ .
\end{aligned}$$

The physical invariants $\mathcal{A}_n$ and $\mathcal{D}_n$ are generic, corresponding respectively to the order 1 (i.e. identity) and order 2 (i.e. the simple-current J) Dynkin diagram symmetries, as we shall see in §1.7. Physically, they are the partition functions of WZW models on $\mathrm{SU}_2(\mathbb{C})$ and $\mathrm{SO}_3(\mathbb{R})$ group manifolds, resp. The exceptionals $\mathcal{E}_6$ and $\mathcal{E}_8$ are best interpreted as due to the $C_{2,1} \supset A_{1,10}$ and $G_{2,1} \supset A_{1,28}$ conformal embeddings (see §1.7; standard notation is to write '$X_{\ell,k}$' for '$X_\ell^{(1)}$ and level k'). The $\mathcal{E}_7$ exceptional is harder to interpret, but can be thought of as the first in an infinite series of exceptionals involving rank-level duality and D_4 triality.

Around Christmas 1985, Zuber wrote Kac about the $A_1^{(1)}$ physical invariant problem, and mentioned the physical invariants he and Itzykson knew at that point (what we now call $\mathcal{A}_\star$ and $\mathcal{D}_{even}$). A few weeks later, Kac wrote back saying he found one more invariant, and jokingly pointed out that it must be indeed quite exceptional as the exponents of E_6 appeared in it. "I must confess that I didn't pay much attention to that last remark (I hardly knew what Coxeter exponents were, at the time!)" [63]. By spring 1986, Cappelli arrived in Paris and got things moving again; together Cappelli-Itzykson-Zuber found $\mathcal{E}_7$, $\mathcal{D}_{odd}$, and then $\mathcal{E}_8$, and struggled to find more. "And it is only in August [1986], during a conversation with Pasquier, in which he was showing me his construction of lattice models based on Dynkin diagrams, that I suddenly remembered this cryptic but crucial! observation of Victor, rushed to the library to find a list of the exponents of the other algebras... and found with the delight that you can imagine that they were matching our list" [63]. Thus the A-D-E pattern to these physical invariants was discovered.

1.6. THE A-D-E META-PATTERN

Before we discuss *meta-patterns* in math, let's introduce the notion of *lattice*[11], a simple geometric structure we'll keep returning to in these notes. The standard reference for lattice theory is [13].

Consider the real vector space $\mathbb{R}^{m,n}$: its vectors look like $\vec{x} = (\vec{x}_+;\vec{x}_-)$ where $\vec{x}_+$ and $\vec{x}_-$ are m- and n-component vectors respectively, and dot products are given by $\vec{x}\cdot\vec{y} = \vec{x}_+\cdot\vec{y}_+ - \vec{x}_-\cdot\vec{y}_-$. The dot products $\vec{x}_\pm\cdot\vec{y}_\pm$ are given by the usual product and sum of components. For example, the familiar Euclidean (positive-definite) space is $\mathbb{R}^n = \mathbb{R}^{n,0}$, while Minkowski space is $\mathbb{R}^{3,1}$.

---

[11] There are many words in math which have several incompatible meanings. For example, there are *vector* fields and *number* fields, and modular *forms* and modular *representations*. 'Lattice' is another of these words. Aside from the geometric meaning we will use, it also refers to a 'partially ordered set'.

Now choose any basis $B = \{\vec{x}_1, \ldots, \vec{x}_{m+n}\}$ in $\mathbb{R}^{m,n}$. So $\mathbb{R}^{m,n} = \mathbb{R}\vec{x}_1 + \cdots + \mathbb{R}\vec{x}_{m+n}$. Define the set $\Lambda(B) := \mathbb{Z}\vec{x}_1 + \cdots + \mathbb{Z}\vec{x}_{m+n}$. This is a *lattice*, and all lattices can be formed in this way[12]. So a lattice is discrete and is closed under sums and integer multiples. For example, $\mathbb{Z}^{m,n}$ is a lattice (take the standard basis in $\mathbb{R}^{m,n}$). A more interesting lattice is the hexagonal lattice (also called A_2), given by the basis $B = \{(\frac{\sqrt{2}}{2}, \frac{\sqrt{6}}{2}), (\sqrt{2}, 0)\}$ of $\mathbb{R}^2$ — try to plot several points. If you wanted to slide a bunch of coins on a table together as tightly as possible, their centres would form this hexagonal lattice. Another important lattice is $II_{1,1} \subset \mathbb{R}^{1,1}$, given by $B = \{(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}; \frac{-1}{\sqrt{2}})\}$; equivalently it can be thought of as the set of all pairs $(a, b) \in \mathbb{Z}^2$ with dot product

$$(a, b) \cdot (c, d) = ad + bc\ . \tag{1.6.1}$$

It is important to note that different choices of basis may or may not result in a different lattice. For a trivial example, consider $B = \{1\}$ and $B' = \{-1\}$ in $\mathbb{R} = \mathbb{R}^{1,0}$: they both give the lattice $\mathbb{Z} = \mathbb{Z}^{1,0}$. Two lattices are called *equivalent* if they only differ by a change-of-basis. E.g. $B = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})\}$ in $\mathbb{R}^2$ yields a lattice equivalent to $\mathbb{Z}^2$.

The *dimension* of the lattice is $m + n$. The lattice is called *positive-definite* if it lies in some $\mathbb{R}^m$ (i.e. $n = 0$). The lattice is called *integral* if all dot products $\vec{x} \cdot \vec{y}$ are integers, for $\vec{x}, \vec{y} \in \Lambda$. A lattice Λ is called *even* if it is integral and in addition all norms $\vec{x} \cdot \vec{x}$ are *even* integers. For example, $\mathbb{Z}^{m,n}$ is integral but not even, while A_2 and $II_{1,1}$ are even. The *dual* Λ^* of a lattice Λ consists of all vectors $\vec{x} \in \mathbb{R}^{m,n}$ such that $\vec{x} \cdot \Lambda \subset \mathbb{Z}$. So a lattice is integral iff $\Lambda \subseteq \Lambda^*$. A lattice is called *self-dual* if $\Lambda = \Lambda^*$. $\mathbb{Z}^{m,n}$ and $II_{1,1}$ are self-dual but A_2 is not.

There are lots of 'meta-patterns' in math, i.e. collections of seemingly different problems which have similar answers. Once one of these meta-patterns is identified it is always helpful to understand what is responsible for it. For example, while I was writing up my PhD thesis I noticed in several places the numbers 1, 2, 3, 4, and 6. For instance $\cos(2\pi r) \in \mathbb{Q}$ for $r \in \mathbb{Q}$ iff the denominator of r is 1, 2, 3, 4, or 6. This pattern was easy to explain: they are precisely those positive integers n with Euler totient $\phi(n) \leq 2$, i.e. there are at most 2 positive numbers less than n coprime[13] to n. The other incidences of these numbers can usually be reduced to this $\phi(n) \leq 2$ property (e.g. the dimension of the number field $\mathbb{Q}[\cos(2\pi\frac{a}{b})]$ (see §1.8) considered as a vector space over $\mathbb{Q}$ will be $\phi(b)/2$).

A more interesting meta-pattern involves the number 24 and its divisors. One sees 24 wherever modular forms naturally appear. For instance, we see it in the critical dimensions in string theory: $24 + 2$ and $8 + 2$. Another example: the dimensions of even self-dual positive-definite lattices must be a multiple of 8 (e.g. the E_8 root lattice defined shortly has dimension 8, while the Leech lattice discussed in §2.4 has dimension 24). The meta-pattern 24 is also understood: the fundamental problem for which it is the answer is the following one. Fix n, and consider the congruence $x^2 \equiv 1 \pmod{n}$. Certainly in order to have a chance of satisfying this, x and n must be coprime. The extreme situation is when *every* number x coprime to n satisfies this congruence:

---

[12] In most presentations a lattice is permitted to have smaller dimension than its ambient space, however that freedom gains no real generality.

[13] We say m,n are *coprime* if any prime p which divides m does not divide n, and vice versa.

$$\gcd(x,n) = 1 \qquad \Longleftrightarrow \qquad x^2 \equiv 1 \pmod{n} \ . \tag{1.6.2}$$

The reader can try to verify the following simple fact: n obeys this extreme situation (1.6.2) iff n divides 24.

What does this congruence property have to do with these other occurrences of 24? Let Λ be an even self-dual positive-definite lattice of dimension n. Then an elementary argument shows that there will exist an n-tuple $\vec{a} = (a_1, \ldots, a_n)$ of odd integers with the property that 8 must divide $\vec{a} \cdot \vec{a} = \sum_i a_i^2$. But $a_i^2 \equiv 1 \pmod 8$, and so we get $8|n$.

A much deeper and still not-completely-understood meta-pattern is called A-D-E (see [1] for a discussion and examples). The name comes from the so-called *simply-laced algebras*, i.e. the simple finite-dimensional Lie algebras whose Dynkin diagrams — see Figure 6 in [59] — contain only single edges (i.e. no arrows). These are the $A_\star$- and $D_\star$-series, along with the E_6, E_7 and E_8 exceptionals. The claim is that many other problems, which don't seem to have anything directly in common with simple Lie algebras, have a solution which falls into this A-D-E pattern (for an object to be meaningfully labelled X_ℓ, some of the data associated to the algebra X_ℓ should reappear in some form in that object). Let's look at some examples.

Consider even positive-definite lattices Λ. The smallest possible nonzero norm in Λ will be 2, and the vectors of norm 2 are special and are called *roots*. The reason they are special is that reflecting through them will always be an automorphism of Λ. That is, the reflection $\vec{u} \mapsto \vec{u} - 2\frac{\vec{u}\cdot\vec{\alpha}}{\vec{\alpha}\cdot\vec{\alpha}}\,\vec{\alpha}$ through $\vec{\alpha} \neq \vec{0}$ won't in general map Λ to itself, unless $\vec{\alpha}$ is a root of Λ. It is important in lattice theory to know the lattices which are spanned by their roots; it turns out these are precisely the orthogonal direct sums of lattices called A_n, D_n, and E_6, E_7 and E_8. They carry those names for a number of reasons. For example, the lattice called X_n will have a basis $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_n\}$ with the property that the matrix $A_{ij} := \vec{\alpha}_i \cdot \vec{\alpha}_j$ is the Cartan matrix (see §2.7) for the Lie algebra X_n! Also, the reflection group generated by reflections in the roots of the lattice X_n will be isomorphic to the Weyl group of the Lie algebra X_n. Finally, to any simple Lie algebra there is canonically associated a lattice called the root lattice; for the simply-laced algebras, these will equal the corresponding lattice of the same name. Incidentally, the root lattices for the non-simply-laced simple algebras will (up to rescalings) be direct sums of the simply-laced root lattices.

We have already met the A_2 lattice: it is the densest packing of circles in the plane. It has long been believed that the obvious pyramidal way to pack oranges is also the densest possible way — the centres of the oranges form the A_3 root lattice. A controversial proof for this famous conjecture has been offered by W.-Y. Hsiang in 1991; in 1998 a new proof by Hale *et al* has been proposed. The densest known packings in dimensions 4,5,6,7,8 are D_4, D_5, E_6, E_7, E_8, resp. E_8 is the smallest even self-dual positive-definite lattice.

A famous A-D-E example is called the McKay[14] correspondence. Consider any finite subgroup G of the Lie group $\mathrm{SU}_2(\mathbb{C})$ (i.e. the 2×2 unitary matrices with determinant 1). For example, there is the cyclic group $\mathbb{Z}_n$ of n elements generated by the matrix

$$M_n = \begin{pmatrix} \exp[2\pi \mathrm{i}/n] & 0 \\ 0 & \exp[-2\pi \mathrm{i}/n] \end{pmatrix}$$

---

[14] He is the same John McKay we will celebrate in section 2.1.

Let R_i be the irreducible representations of G. For instance, for $\mathbb{Z}_n$, there are precisely n of these, all 1-dimensional, given by sending the generator M_n to $\exp[2k\pi\mathrm{i}/n]$ for each $k = 1, 2, \ldots, n$. Now consider the tensor product $G \otimes R_i$, where we interpret $G \subset \mathrm{SU}_2(\mathbb{C})$ here as a 2-dimensional representation. We can decompose that product into a direct sum $\oplus_j m_{ij} R_j$ of irreducibles (the m_{ij} here are multiplicities). Now create a graph with one node for each R_i, and with the ith and jth nodes ($i \neq j$) connected with precisely m_{ij} directed edges $i \to j$. If $m_{ij} = m_{ji}$, we agree to erase the double arrows from the m_{ij} edges. Then McKay observed that the graph of any G will be a distinct extended Dynkin diagram of A-D-E type! For instance, the cyclic group with n elements corresponds to the extended graph of A_{n-1}.

How was McKay led to his remarkable correspondence? He knew that the sum of the 'marks' $a_i = 1, 2, 3, 4, 5, 6, 4, 2, 3$ associated to each node of the extended E_8 Dynkin diagram equaled 30, the Coxeter number of E_8. So what did their *squares* add to? 120, which he recognised as the cardinality of one of the exceptional finite subgroups of $\mathrm{SU}_2(\mathbb{C})$, and that got him thinking...

Another famous example of A-D-E, due to Arnol'd, are the 'simple critical points' of smooth complex-valued functions f, on e.g. $\mathbb{C}^3$. For example, both $x^2 + y^2 + z^{n+1}$ and $x^2 + y^3 + z^5$ have singularities at $(0, 0, 0)$ (i.e. their first partial derivatives all vanish there), and they are assigned to A_n and E_8, respectively. The $\mathrm{SU}_2(\mathbb{C})$ subgroups can be related to singularities as follows. The group $\mathrm{SU}_2(\mathbb{C})$ acts on $\mathbb{C}^2$ in the obvious way (matrix multiplication). If G is a discrete subgroup, then consider the (ring of) polynomials in 2 variables w_1, w_2 invariant under G. It turns out it will have 3 generators $x(w_1, w_2)$, $y(w_1, w_2)$, $z(w_1, w_2)$, which are connected by 1 polynomial relation (syzygy). For instance, take G to be the cyclic group $\mathbb{Z}_n$, then we're interested in polynomials $p(w_1, w_2)$ invariant under $w_1 \mapsto \exp[2\pi\mathrm{i}/n] w_1$, $w_2 \mapsto \exp[-2\pi\mathrm{i}/n] w_2$. Any such invariant $p(w_1, w_2)$ is clearly generated by (i.e. can be written as a polynomial in) $w_1 w_2$, w_1^n and w_2^n. Choosing instead the generators $x = \frac{w_1^n - w_2^n}{2}$, $y = \mathrm{i}\,\frac{w_1^n + w_2^n}{2}$, $z = w_1 w_2$, we get the syzygy $z^n = -(x^2 + y^2)$. For any G, generators x, y, z can always be found so that the syzygy will be one of the polynomials associated to a simple singularity, and in fact will give the equation of the algebraic surface $\mathbb{C}^2/G$ as a 2-dimensional complex surface in $\mathbb{C}^3$ (e.g. the complex surfaces $\mathbb{C}^2/\mathbb{Z}_n$ and $\{(x, y, z) \in \mathbb{C}^3 \,|\, x^2 + y^2 + z^n = 0\}$ are equivalent).

Arguably the first A-D-E classification goes back to Theaetetus, around 400 B.C. He classified the regular solids. For instance the tetrahedron can be associated to E_6 while the cube is matched with E_7. This A-D-E is only partial, as there are no regular solids assigned to the A-series, and to get the D-series one must look at 'degenerate regular solids'.

The closest thing to an explanation of the A-D-E meta-pattern would seem to be the notion of 'additive assignments' on graphs (which is a picturesque way of describing the corresponding eigenvalue problem). Consider any graph $\mathcal{G}$ with undirected edges, and none of the edges run from a node to itself. We can also assume without loss of generality that $\mathcal{G}$ is connected. Assign a positive number a_i to each node. If this assignment has the property that for each i, $2a_i = \sum a_j$ where the sum is over all nodes j adjacent to i (counting multiplicities of edges), then we call it 'additive'. For instance, for the graph ○═○, the assignment $a_1 = 1 = a_2$ is additive, but the assignment $a_1 = 1, a_2 = 2$ is not. The question is, which graphs have an additive assignment? The answer is: precisely the extended Dynkin diagrams of A-D-E type! And their additive assignments are unique (up to constant proportionality) and are given by the *marks* a_i of the algebra (see e.g. the Table on p.54 of [38]). For example the extended A_n graph consists of $n+1$ nodes arranged in a circle, and its marks a_i all equal 1.

What do additive assignments have to do with the other A-D-E classifications? Consider a finite subgroup G of $\mathrm{SU}_2(\mathbb{C})$. Take the dimension of the equation $G \otimes R_i = \oplus_j m_{ij} R_j$: we get $2d_i = \sum_j m_{ij} d_j$ where $d_j = \dim(R_j)$. Hence the dimensions of the irreducible representations define an additive assignment for each of McKay's graphs, and hence those graphs must be of A-D-E type (provided we know $m_{ij} = m_{ji}$).

As Cappelli-Itzykson-Zuber observed, the physical invariants for $A_1^{(1)}$ also realise the A-D-E pattern, in the following sense. The *Coxeter number* h of the name $\mathcal{X}_\ell$ (i.e. the sum $\sum_i a_i$ of the marks) equals $k+2$, and the *exponents* m_i of X_ℓ equal those $a \in P_+^k$ for which $M_{aa} \neq 0$ (for the simply-laced algebras, the m_i are defined by writing the eigenvalues of the corresponding Cartan matrix (see §2.7) as $4\sin^2(\frac{\pi m_i}{2h})$ — the m_i are integers and the smallest is always 1). Probably what first led Kac to his observation about the E_6 exponents was that $k+2$ (this is how k enters most formulas) for his exceptional equalled the Coxeter number 12 for E_6. More recently, the operator algebraists Ocneanu [48] and independently Böckenhauer-Evans [4] found an A-D-E interpretation for the off-diagonal entries M_{ab} of the $A_1^{(1)}$ physical invariants, using subfactor theory.

We are not claiming that this $A_1^{(1)}$ classification is 'equivalent' to any other A-D-E one — that would miss the point of meta-patterns. What we really want to do is to identify some critical combinatorial part of an $A_1^{(1)}$ proof with critical parts in other A-D-E classifications — this is what we did with the other meta-patterns. A considerably simplified proof of the $A_1^{(1)}$ classification is now available [29], so hopefully this task will now be easier.

There has been some progress at understanding this $A_1^{(1)}$ A-D-E. Nahm [46] constructed the invariant $\mathcal{X}_\ell$ in terms of the compact simply-connected Lie group of type X_ℓ, and in this way could interpret the $k+2=h$ and $M_{m_i m_i} \neq 0$ coincidences. A very general explanation for A-D-E has been suggested by Ocneanu [48] using his theory of path algebras on graphs; although his work has never been published, others are now rediscovering (and publishing!) similar work (see e.g. [4]). Nevertheless, the A-D-E in CFT remains almost as mysterious now as it did a dozen years ago — for example it still isn't clear how it directly relates to additive assignments.

There are 4 other claims for A-D-E classifications of families of RCFT physical invariants, and all of them inherit their (approximate) A-D-E pattern from the more fundamental $A_1^{(1)}$ one. One is the $c<1$ minimal models, also proven in [8], and another is the $N=1$ superconformal minimal models, proved by Cappelli (1987). In both cases the physical invariants are parametrised by pairs of A-D-E diagrams. The list of known $c=1$ RCFTs also looks like A-D-E (two series parametrised by $\mathbb{Q}_+$, and three exceptionals), but the completeness of that list has never been rigourously established.

The fourth classification often quoted as A-D-E, is the $N=2$ superconformal minimal models. Their classification was done by Gannon (1997). The connection here with A-D-E turns out to be rather weak: e.g. 20, 30, and 24 distinct invariants would have an equal right to be called $\mathcal{E}_6$, $\mathcal{E}_7$, and $\mathcal{E}_8$ respectively. It would appear that the frequent claims that the $N=2$ minimal models fall into an A-D-E pattern are rather dubious.

Hanany-He [35] suggest that the $A_1^{(1)}$ A-D-E pattern can be related to subgroups $G \subset \mathrm{SU}_2(\mathbb{C})$ by orbifolding 4-dimensional $N=4$ supersymmetric gauge theory by G, resulting in an $N=2$ superCFT whose 'matter matrix' can be read off from the Dynkin diagram corresponding to G. The same game can be played with finite subgroups of $\mathrm{SU}_3(\mathbb{C})$, resulting in $N=1$ superCFTs whose matter matrices correspond to graphs very reminiscent of the 'fusion graphs' of Di Francesco-Petkova-Zuber (see e.g. [62]) corresponding to $A_2^{(1)}$ physical invariants. [35] use this to conjecture a McKay-type correspondence between singularities of type $\mathbb{C}^n/G$, for $G \subset \mathrm{SU}_n(\mathbb{C})$, and the physical invariants of $A_{n-1}^{(1)}$. This in their view would be the form A-D-E takes for higher rank physical invariants. Their actual conjecture though is still somewhat too vague.

For a final example of meta-pattern, consider 'modular function' (see §2.3). After all, they appear in a surprising variety of places and disguises. Maybe we shouldn't regard their ubiquity as fortuitous, instead perhaps there's a deeper common 'situation' which is the source for that ubiquity. Just as 'symmetry' yields 'group', or 'rain-followed-by-heat' breeds mosquitos. Math is not above metaphysics; like any area it grows by asking questions, and changing your perspective — even to a metaphysical one — should suggest new questions.

1.7. SIMPLE-CURRENTS AND CHARGE-CONJUGATION

The key properties[15] of the matrix S are that it's unitary and symmetric (so M in §1.5 equals SMS^*),

$$S_{0\mu} > 0 \text{ for all } \mu \in P_+^k \ , \tag{1.7.1}$$

and that the numbers $N_{\lambda\mu}^{\nu}$ defined by Verlinde's formula (1.4.4) are nonnegative integers. These are obeyed by the matrix S in any (unitary) RCFT. From these basic properties, we will obtain here some elementary consequences which have important applications.

But first, let's make an observation which isn't difficult to prove, but doesn't appear to be generally known.

Verlinde's formula looks strange, but it is quite generic,

and we can see it throughout math and mathematical physics. Consider the following.

Let $\mathfrak{A}$ be a commutative associative algebra, over $\mathbb{R}$ say. Suppose $\mathfrak{A}$ has a finite basis Φ (over $\mathbb{R}$) containing the unit 1. Define the 'structure constants' $N_{ab}^c \in \mathbb{R}$, for $a,b,c \in \Phi$, by $ab = \sum_{c\in\Phi} N_{ab}^c c$. Suppose there is an algebra homomorphism $*$ (so $*$ is linear, and $(xy)^* = x^* y^*$) which permutes the basis vectors (so $\Phi^* = \Phi$), and we have the relation $N_{ab}^1 = \delta_{b,a^*}$. We call any such algebra $\mathfrak{A}$ a *fusion algebra.* Then any fusion algebra will necessarily have a unitary matrix S with $S_{1a} > 0$ and with the structure constants given by Verlinde's formula. Algebraically, the relation $S = S^t$ holds if $\mathfrak{A}$ is 'self-dual' in a certain natural sense.

---

[15] A good exercise for the reader is to prove that if S is unitary and symmetric, and obeys (1.7.1), then there will be at most finitely many physical invariants M for that S,T.

Define the 'fusion matrices' N_λ by $(N_\lambda)_{\mu\nu} = N_{\lambda\mu}^\nu$. Then Verlinde's formula says that the μth column $S_{\uparrow\mu}$ of S is an eigenvector of each fusion matrix N_λ, with eigenvalue $\frac{S_{\lambda\mu}}{S_{0\mu}}$.

Useful Fact. If v is a simultaneous eigenvector of each fusion matrix N_λ, then there exists a constant $c \in \mathbb{C}$ and a $\lambda \in P_+^k$ such that $v = c\,S_{\uparrow\lambda}$.

For one consequence, take the complex conjugate of the eigenvector equation $N_\lambda\, S_{\uparrow\mu} = \frac{S_{\lambda\mu}}{S_{0\mu}}\, S_{\uparrow\mu}$: we get that the vector $S_{\uparrow\mu}^*$ is a simultaneous eigenvector of all N_λ, and hence must equal $c\,S_{\uparrow\gamma}$ for some number c and weight $\gamma \in P_+^k$, both depending on μ. Write $\gamma = C\mu$; then C defines a permutation of P_+^k. The reader can verify that unitarity of S forces $|c| = 1$, while (1.7.1) forces $c > 0$. Thus $c = 1$ and we obtain the formula

$$S_{\lambda\mu}^* = S_{\lambda,C\mu} = S_{C\lambda,\mu}\ . \tag{1.7.2}$$

Also, unitarity and symmetry of S forces $C = S^2$, while conjugating twice shows $C^2 = id.$ C is an important matrix in RCFT, and is called *charge-conjugation*. When $C = id.$, then the matrix S is real.

Note that (1.7.1) now implies $C0 = 0$. Also $CT = TC$. Hence $M = C$ will always define a physical invariant, and if M is any other physical invariant, the matrix product $MC = CM$ will define another physical invariant. Also, $N_{C\lambda,C\mu}^{C\nu} = N_{\lambda\mu}^\nu$ and $(N_\lambda)^t = CN_\lambda = N_\lambda C = N_{C\lambda}$.

For the WZW (=affine) case, C has a special meaning: $C\lambda$ is the highest-weight 'contragredient' to λ. C corresponds to an order 2 (or 1) symmetry of the (unextended) Dynkin diagram. For example, for $A_\ell^{(1)}$, we have $C(\lambda_0, \lambda_1, \ldots, \lambda_{\ell-1}, \lambda_\ell) = (\lambda_0, \lambda_\ell, \lambda_{\ell-1}, \ldots, \lambda_1)$. For $A_1^{(1)}$ then, $C = id.$, which can also be read off from (1.4.3).

The algebras $D_{even}^{(1)}$ all have at least one nontrivial symmetry of the (unextended) Dynkin diagram which isn't the charge-conjugation. The most interesting example is $D_4^{(1)}$, which has 5 of these. By a *conjugation*, we will mean any symmetry of the unextended Dynkin diagram.

To go much further, we need a fascinating tool called *Perron-Frobenius theory* — a collection of results concerning the eigenvalues and eigenvectors of nonnegative matrices (i.e. matrices in which every entry is a nonnegative real number). Whenever you have such matrices in your problem, and it is natural to multiply them, then there is a good chance Perron-Frobenius theory will tell you something interesting. The basic result here is that if A is a nonnegative matrix, then there will be a nonnegative eigenvector $x \geq 0$ with eigenvalue $\rho \geq 0$, such that if λ is any other eigenvalue of A, then $|\lambda| \leq \rho$. There are lots of other results (see e.g. [44]), e.g. ρ must be at least as large as any diagonal entry of A, and there must be a row-sum of A no bigger than ρ, and another row-sum no smaller than ρ.

For instance, consider

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\ , \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\ .$$

Perron-Frobenius eigenvectors for A and B are $\begin{pmatrix}1\\1\\1\end{pmatrix}$ and $\begin{pmatrix}2\\1\\1\end{pmatrix}$, with eigenvalues 3 and 2 resp. The other eigenvalue of A is 0 (multiplicity 2), while those of B are 0 and -1.

Fusion matrices N_λ are nonnegative, and it is indeed natural to multiply them:

$$N_\lambda N_\mu = \sum_{\nu\in P_+^k} N_{\lambda\mu}^\nu N_\nu \ .$$

So we can expect Perron-Frobenius to tell us something interesting. This is the case, and we obtain the curious-looking inequalities

$$S_{\lambda 0}\, S_{0\mu} \geq |S_{\lambda\mu}|\, S_{00} \ . \tag{1.7.3}$$

Squaring both sides, summing over μ and using unitarity, we get that $S_{\lambda 0} \geq S_{00}$. In other words, the ratio $\frac{S_{\lambda 0}}{S_{00}}$, called the *quantum-dimension* of λ, will necessarily be ≥ 1.

The term 'quantum-dimension' comes from quantum groups, where $\frac{S_{\lambda 0}}{S_{00}}$ is the quantum-dimension of the module labelled by λ of the quantum group $U_q(X_\ell)$.

The borderline case then [16] is when a quantum-dimension *equals* 1. Any such weight is called a *simple-current*. The theory of simple-currents was developed most extensively by Schellekens and collaborators (see e.g. [50]). The simple-currents for the affine algebras were classified by J. Fuchs (1991), and the result is that (with one unimportant exception: $E_8^{(1)}$ at level 2) they all correspond to symmetries of the extended Dynkin diagrams. In particular, applying any such symmetry to the vacuum $0 = (k, 0, \ldots, 0)$ gives the list of simple-currents. For instance, the $\ell + 1$ weights of the form $(0, \ldots, k, \ldots, 0)$ (k in the ith spot) are the simple-currents for $A_\ell^{(1)}$. There are 2 simple-currents for $B_\ell^{(1)}$, $C_\ell^{(1)}$ and $E_7^{(1)}$, 3 for $E_6^{(1)}$, and 4 for $D_\ell^{(1)}$. Simple-currents play a large role in RCFT, as we shall see.

Let j be any simple-current. Then (1.7.3) becomes $S_{0\mu} \geq |S_{j\mu}|$ for all μ, so unitarity forces $S_{0\mu} = |S_{j\mu}|$, that is

$$S_{j\mu} = \exp[2\pi \mathrm{i}\, Q_j(\mu)]\, S_{0\mu} \qquad \forall \mu \in P_+^k \tag{1.7.4}$$

for some rational numbers $0 \leq Q_j(\mu) < 1$. Hence by diagonalising, we get $N_j N_{Cj} = I$. But the inverse of a nonnegative matrix A is itself nonnegative, only if A is a 'generalised permutation matrix', i.e. a permutation matrix except the 1's can be replaced by any positive numbers. But N_j and N_{Cj} are also integral, and so they must in fact be permutation matrices. Write $(N_j)_{\lambda\mu} = \delta_{\mu, J\lambda}$ for some permutation J of P_+^k. So $j = J0$. Then

$$\delta_{\lambda,\mu} = N_{j,\lambda}^{J\mu} = \sum_\nu \exp[2\pi \mathrm{i}\, Q_j(\nu)]\, S_{\lambda\nu}\, S_{J\mu,\nu}^* \ ,$$

---

[16] This seems to be a standard trick in math: when some sort of bound is established, look at the extremal cases which realise that bound. If your bound is a good one, it should be possible to say something about those extremal cases, and having something to say is always of paramount importance. This trick is used for instance in the definition of 24 last section, and the definition of normal subgroup in section 2.2.

so taking absolute values and using the triangle inequality and unitarity of S, we find that (1.7.4) generalises:

$$S_{J\lambda,\mu} = \exp[2\pi \mathrm{i}\, Q_j(\mu)]\, S_{\lambda\mu} \ . \tag{1.7.5}$$

The simple-currents form a finite abelian group, corresponding to the composition of the permutations J. For any simple-currents J, J', we get the symmetry $N_{J\lambda,J'\mu}^{\nu} = N_{\lambda\mu}^{JJ'\nu}$. The $\mathbb{Q}/\mathbb{Z}$-valued functions Q_j define gradings on the fusion rings, and conversely any grading corresponds to a simple-current in this way.

For example, the simple-current $j = (0, k)$ of $A_1^{(1)}$ at level k corresponds to $Q_j(\lambda) = \lambda_1/2$ and the permutation $J\lambda = (\lambda_1, \lambda_0)$. We can see this directly from (1.4.3). For $A_2^{(1)}$ level k, there are 2 nontrivial simple-currents, $(0, k, 0)$ and $(0, 0, k)$. The first of these corresponds to triality $Q(\lambda) = (\lambda_1 + 2\lambda_2)/3$ and $\lambda \mapsto (\lambda_2, \lambda_0, \lambda_1)$, while the second to $(2\lambda_1 + \lambda_2)/3$ and $\lambda \mapsto (\lambda_1, \lambda_2, \lambda_0)$. Similar statements hold for all affine algebras: e.g. for $B_\ell^{(1)}$ level k, the nontrivial simple-current has $Q_j(\lambda) = \lambda_\ell/2$ and $J\lambda = (\lambda_1, \lambda_0, \lambda_2, \ldots, \lambda_\ell)$.

One of the applications of simple-currents is that physical invariants can be built from them in generic ways. These physical invariants all obey the selection rule

$$M_{\lambda\mu} \neq 0 \quad \Longrightarrow \quad \mu = J\lambda \quad \text{for some simple}-\text{current } J = J(\lambda, \mu) \ . \tag{1.7.6}$$

We will call any such physical invariant M a *simple-current invariant.* A special case is the $\mathcal{D}_{\frac{k}{2}+2}$ physical invariant for $A_1^{(1)}$ at even level k. Up to a fairly mild assumption, all simple-current invariants have been classified for any RCFT by Schellekens and collaborators; given that assumption, they can all be constructed by generic methods. The basic construction is due to Bernard [3], though it has been generalised by others. In the WZW case, all simple-current invariants (except some for $D_\ell^{(1)}$) correspond to strings on nonsimply-connected Lie groups.

By a *generic physical invariant* of $X_\ell^{(1)}$ we mean one of the form $M = C'M'C''$ where C', C'' are (charge-)conjugations, and M' is a simple-current invariant. In other words, M is constructed in generic ways from symmetries of the extended Dynkin diagram of X_ℓ. Any other M are called *exceptional.*

All known results point to the validity of the following guess:

Conjecture. Choose a simple algebra X_ℓ. Then for all sufficiently large k, all physical invariants of $X_\ell^{(1)}$ at level k will be generic.

In other words, any given $X_\ell^{(1)}$ will have only finitely many exceptionals. For instance, for $A_1^{(1)}$ and $A_2^{(1)}$ at any $k > 28$ and $k > 21$ resp., all physical invariants are generic. For $C_2^{(1)}$ and $G_2^{(1)}$, $k > 12$ and $k > 4$ resp. should work.

The richest source of exceptionals are *conformal embeddings.* In some cases the affine representations L_λ for some algebra $X_\ell^{(1)}$ (necessarily at level 1) can be decomposed into *finite* direct sums of representations of some affine subalgebra $Y_m^{(1)}$ (at some level k). In this case, a physical invariant for $X_\ell^{(1)}$ level 1 will yield a physical invariant for $Y_m^{(1)}$ level k, obtained by replacing every $X_\ell^{(1)}$ level 1 character χ_λ by the appropriate finite sum of

$Y_m^{(1)}$ level k characters. An example will demonstrate this simple idea: $A_1^{(1)}$ level 28 is a conformal subalgebra of $G_2^{(1)}$ level 1, and we have the character decompositions

$$\begin{aligned}\chi_{(1,0,0)} &= \chi'_{(28,0)} + \chi'_{(18,10)} + \chi'_{(10,18)} + \chi'_{(0,28)} \\ \chi_{(0,0,1)} &= \chi'_{(22,6)} + \chi'_{(16,12)} + \chi'_{(12,16)} + \chi'_{(6,22)} \ .\end{aligned}$$

Thus the unique level 1 $G_2^{(1)}$ physical invariant $|\chi_0|^2 + |\chi_{(0,0,1)}|^2$ yields what we call the $\mathcal{E}_8$ physical invariant of $A_1^{(1)}$. All level 1 physical invariants are known, as are all conformal embeddings and the corresponding character decompositions (branching rules).

1.8. Galois Theory

Evariste Galois was a brilliantly original French mathematician. Born shortly before Napoleon's ill-fated invasion of Russia, he died shortly before the ill-fated 1832 uprising in Paris. His last words: "Don't cry, I need all my courage to die at 20".

Galois grew up in a time and place confused and excited by revolution. He was known to say "if I were only sure that a body would be enough to incite the people to revolt, I would offer mine". On May 2 1832, after frustration over failure in love and failure to convince the Paris math establishment of the depth of his ideas, he made his decision. A duel was arranged with a friend, but only his friend's gun would be loaded. Galois died the day after a bullet perforated his intestine. At his funeral it was discovered that a famous general had also just died, and the revolutionaries decided to use the general's death rather than Galois' as a pretext for an armed uprising. A few days later the streets of Paris were blocked by barricades, but not because of Galois' sacrifice: his death had been pointless [56].

Galois theory in its most general form is the study of relations between objects defined implicitly by some conditions. For example, the objects could be the solutions to a given differential equation. In the incarnation of Galois we are interested in here, the objects are numbers, namely the zeros of certain polynomials. We will sketch this theory below, but see e.g. the article by Stark in [60] for more details.

Gauss seems to have been the first to show that 'weird' (complex) numbers could tell us about the integers. For instance, suppose we are interested in the equation $n = a^2 + b^2$. Consider $5 = 2^2 + 1^2$. We can write this as $5 = (2 + \mathrm{i})(2 - \mathrm{i})$, so we are led to consider complex numbers of the form $a + b\mathrm{i}$, for $a, b \in \mathbb{Z}$. These are now called 'Gaussian integers'. Suppose we know the following theorem:

Fact. Let $p \in \mathbb{Z}$ be any prime number. Then p factorises over the Gaussian integers iff $p = 2$ or $p \equiv 1 \pmod 4$.

By 'factorise' there, we mean $p = zw$ where neither z nor w is a 'unit': ± 1, $\pm\mathrm{i}$.

Now suppose p is a prime, $= 2$ or $\equiv 1 \pmod 4$, and we write $p = (a + b\mathrm{i})(c + d\mathrm{i})$. Then $p^2 = (a^2 + b^2)(c^2 + d^2)$, so $a^2 + b^2 = c^2 + d^2 = p$. Conversely, suppose $p = a^2 + b^2$, then $p = (a + b\mathrm{i})(a - b\mathrm{i})$. Thus:

Consequence. [17] Let $p \in \mathbb{Z}$ be any prime number. Then $p = a^2 + b^2$ for $a, b \in \mathbb{Z}$ iff $p = 2$ or $p \equiv 1 \pmod 4$.

Now we can answer the question: can a given n be written as a sum of 2 squares $n = a^2 + b^2$? Write out the prime decomposition $n = \prod p^{a_p}$. Then $n = a^2 + b^2$ has a solution iff a_p is *even* for every $p \equiv 3 \pmod 4$. For instance $60 = 2^2 \cdot 3^1 \cdot 5^1$ cannot be written as the sum of 2 squares, but $90 = 2^1 \cdot 3^2 \cdot 5^1$ can. We can also find (and count) all solutions: e.g. $90 = 2 \cdot 3^2 \cdot 5 = \{(1+\mathrm{i})3(1+2\mathrm{i})\}\{(1-\mathrm{i})3(1-2\mathrm{i})\}$, giving $90 = (-3)^2 + 9^2$.

This problem should give the reader a small appreciation for the power of using nonintegers to study integers. Nonintegers often lurk in the shadows, secretly watching their more arrogant brethren the integers strut. One of the consequences of their presence can be the existence of certain 'Galois' symmetries. Such happens in RCFT, as we will show below.

Look at complex conjugation: $(wz)^* = w^* z^*$ and $(w+z)^* = w^* + z^*$. Also, $r^* = r$ for any $r \in \mathbb{R}$. So we can say that $*$ is a structure-preserving map $\mathbb{C} \to \mathbb{C}$ (called an *automorphism* of $\mathbb{C}$) fixing the reals. We will write this $* \in \mathrm{Gal}(\mathbb{C}/\mathbb{R})$. '$\mathrm{Gal}(\mathbb{C}/\mathbb{R})$' is the Galois group of $\mathbb{C}$ over $\mathbb{R}$; it turns out to contain only $*$ and the identity.

A way of thinking about the automorphism $*$ is that it says that, as far as the real numbers are concerned, i and $-$i are identical twins.

Let $\mathbb{F}$ be any *field* containing $\mathbb{Q}$ (we defined 'field' in §1.2). The *Galois group* $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ then will be the set of all automorphisms=symmetries of $\mathbb{F}$ which fix all rationals.

For example, take $\mathbb{F}$ to be the set of all numbers of the form $a + b\sqrt{5}$, where $a, b \in \mathbb{Q}$. Then $\mathbb{F}$ will be a field, which is commonly denoted $\mathbb{Q}[\sqrt{5}]$ because it is generated by $\mathbb{Q}$ and $\sqrt{5}$. Let's try to find its Galois group. Let $\sigma \in \mathrm{Gal}(\mathbb{F}/\mathbb{Q})$. Then $\sigma(a + b\sqrt{5}) = \sigma(a) + \sigma(b)\sigma(\sqrt{5}) = a + b\sigma(\sqrt{5})$, so once we know what σ does to $\sqrt{5}$, we know everything about σ. But $5 = \sigma(5) = \sigma(\sqrt{5}^2) = (\sigma(\sqrt{5}))^2$, so $\sigma(\sqrt{5}) = \pm\sqrt{5}$ and there are precisely 2 possible Galois automorphisms here (one is the identity). As far as $\mathbb{Q}$ is concerned, $\pm\sqrt{5}$ are interchangeable: it cannot see the difference.

For a more important example, consider the *cyclotomic field* $\mathbb{F} = \mathbb{Q}[\xi_n]$, where $\xi_n := \exp[2\pi\mathrm{i}/n]$ is an nth root of 1. So $\mathbb{Q}[\xi_n]$ consists of all complex numbers which can be expressed as polynomials $a_m\xi_n^m + a_{m-1}\xi_n^{m-1} + \cdots + a_0$ in ξ_n with rational coefficients a_i. Once again, to find the Galois group $\mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$, it is enough to see what an automorphism σ does to the generator ξ_n. Since $\xi_n^n = 1$, we see that it must send it to another nth root of 1, ξ_n^ℓ say; in fact it is easy to see that $\sigma(\xi_n)$ must be another 'primitive' nth root of 1, i.e. ℓ must be coprime to n. So $\mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$ will be isomorphic to the multiplicative group $\mathbb{Z}_n^\times$ of numbers between 1 and n coprime to n. The rationals can't see any difference between the primitive nth roots of 1 — for instance $\mathbb{Q}$ can't tell that $\xi_n^{\pm 1}$ are 'closer to 1' than the other primitive roots. So any $\sigma \in \mathrm{Gal}(\mathbb{Q}[\xi_n]/\mathbb{Q})$ will correspond to some $\ell \in \mathbb{Z}_n^\times$, and to see what σ does to some $z \in \mathbb{Q}[\xi_n]$ what we do is write z as a polynomial in ξ_n and then replace each occurrence of ξ_n with ξ_n^ℓ. For example,

$$\sigma\big(\cos(2\pi a/n)\big) = \sigma\left(\frac{\xi_n^a + \xi_n^{-a}}{2}\right) = \frac{\xi_n^{a\ell} + \xi_n^{-a\ell}}{2} = \cos(2\pi a\ell/n) \ .$$

---

[17] This result was first stated by Fermat in one of his infamous margin notes (another is discussed in Section 2.3), and was finally proved a century later by Euler. A remarkable 1-line proof was found by Zagier [61].

So to summarise, Galois automorphisms are a massive generalisation of the idea of complex conjugation. If in your problem complex conjugation seems interesting, then there is a good chance more general Galois automorphisms will play an interesting role. This is what happens in RCFT, as we now show.

Fact. [14] Suppose S is unitary and symmetric and each $S_{0a} > 0$.

(a) If in addition the numbers N_{ab}^c given by Verlinde's formula (1.4.4) are rational, then the entries S_{ab} of S must lie in a cyclotomic field.

(b) The numbers N_{ab}^c will be rational iff for any $\sigma \in \mathrm{Gal}(\mathbb{Q}[S]/\mathbb{Q})$, there is a permutation $a \mapsto a^\sigma$, and a choice of signs $\epsilon_\sigma(a) \in \{\pm 1\}$, such that

$$\sigma(S_{ab}) = \epsilon_\sigma(a)\, S_{a^\sigma,b} = \epsilon_\sigma(b)\, S_{a,b^\sigma} \ . \tag{1.8.1}$$

'$\mathbb{Q}[S]$' in part (b) denotes the field generated by $\mathbb{Q}$ and all matrix entries S_{ab}. The argument follows the one given for the charge-conjugation C at the beginning of the last section. The kinds of complex numbers which lie in cyclotomic fields are $\sin(\pi r)$, $\cos(\pi r)$, $\sqrt{r}$ and $r\mathrm{i}$ for any $r \in \mathbb{Q}$. Almost all complex numbers fail to lie in any cyclotomic field: e.g. generic cube roots, 4th roots, ..., of rationals, as well as transcendental numbers like e, π and e^π.

Of course the affine algebras satisfy the conditions of the Fact, as does more generally the modular matrix S for any unitary RCFT, and so these will possess the Galois action. For the affine algebras this action has a geometric interpretation in terms of multiplying weights by an integer ℓ and applying Weyl group elements — see [14] for a description.

This Fact is useful in both directions: as a way of testing whether a conjectured matrix S has a chance of producing the integral fusions we want it to yield; and more importantly as a source of a symmetry of the RCFT which generalises charge-conjugation. Any statement about charge-conjugation seems to have an analogue for any of these Galois symmetries, although it is usually more complicated.

As an example, consider $A_1^{(1)}$: (1.4.3) shows explicitly that $S_{\lambda\mu}$ lies in the cyclotomic field $\mathbb{Q}[\xi_{4(k+2)}]$. Write $\{x\}$ for the number congruent to x mod $2(k+2)$ satisfying $0 \le \{x\} < 2(k+2)$. Choose any Galois automorphism σ, and let $\ell \in \mathbb{Z}_{4(k+2)}^\times$ be the corresponding integer. Then if $\{\ell(a+1)\} < k+2$, we will have $a^\sigma = \{\ell(a+1)\} - 1$, while if $\{\ell(a+1)\} > k+2$, we'll have $a^\sigma = 2(k+2) - \{\ell(a+1)\} - 1$. The sign $\epsilon_\sigma(a)$ will depend on a contribution from $\sqrt{\frac{2}{k+2}}$ (which for most purposes can be ignored), as well as the sign $+1$ or -1, resp., depending on whether or not $\{\ell(a+1)\} < k+2$.

Consider specifically $k = 10$, and the Galois automorphism σ_5 corresponding to $\ell = 5$. Then the permutation is $0 \leftrightarrow (6,4)$, $(9,1) \leftrightarrow (1,9)$, $(8,2) \leftrightarrow (2,8)$, $(4,6) \leftrightarrow (0,10)$, while $(7,3)$ and (5,5) are fixed.

This Galois symmetry has been used to find certain exceptional physical invariants, but its greatest use so far is as a powerful selection rule we will describe next section.

In this final section we include some of the basic tools belonging to the 'modern' classifications of physical invariants, and we give a flavour of their proofs. We will state them for the $A_1^{(1)}$ level k problem given above, but everything generalises without effort. See [29] and references therein for more details. Recall the matrices S, T in (1.4.3).

First note that commutation of M with T implies the selection rule

$$M_{\lambda\mu} \neq 0 \quad \Longrightarrow \quad (\lambda_1+1)^2 \equiv (\mu_1+1)^2 \pmod{4(k+2)} \ . \tag{1.9.1}$$

It is much harder to squeeze information out of the commutation with S, but the resulting information turns out surprisingly to be much more useful. In fact, commutation with S is almost incompatible with the constraint $M_{\lambda\mu} \in \mathbb{Z}_{\geq}$.

Note that the vacuum $0 \in P_+^k$ is both physically and mathematically special; our strategy will be to find all possible 0th rows and columns of M, and then for each of these possibilities to find the remaining entries of M.

The easiest result follows by evaluating $MS = SM$ at $(0, \lambda)$ for any $\lambda \in P_+^k$:

$$\sum_{\mu \in P_+^k} M_{0\mu}\, S_{\mu\lambda} \geq 0 \ , \tag{1.9.2}$$

with equality iff the λth column of M is identically 0. (1.9.2) has two uses: it severely constrains the values of $M_{0\mu}$ (similarly $M_{\mu 0}$), and it says precisely which columns (and rows) are nonzero.

Next, let's apply the triangle inequality to sums involving (1.7.5). Choose any $i, j \in \{0,1\}$. Then

$$M_{J^i 0, J^j 0} = \sum_{\lambda,\mu} (-1)^{\lambda_1 i}\, S_{0\lambda}\, M_{\lambda\mu}\, (-1)^{\mu_1 j}\, S_{0\mu} \ .$$

Taking absolute values, we obtain

$$M_{J^i 0, J^j 0} \leq \sum_{\lambda,\mu} S_{0\lambda}\, M_{\lambda\mu}\, S_{0\mu} = M_{00} = 1 \ .$$

Thus $M_{J^i 0, J^j 0}$ can equal only 0 or 1. If it equals 1, then we obtain the selection rule:

$$\lambda_1 i \equiv \mu_1 j \pmod 2 \text{ whenever } M_{\lambda\mu} \neq 0 \ ;$$

this implies the symmetry $M_{J^i\lambda, J^j\mu} = M_{\lambda\mu}$ for all $\lambda, \mu \in P_+^k$. We can see both of these in the list of physical invariants for $A_1^{(1)}$ level k. This explains a lot of the properties of those invariants. For instance, try to use this selection rule to explain why no χ_{odd} appears in the exceptional called $\mathcal{E}_8$.

Our M is nonnegative, and although multiplying M's may not give us back a physical invariant, it will give us a matrix commuting with S and T. In other words, the commutant is much more than merely a vector space, it is in fact an algebra. Thus we should expect Perron-Frobenius to tell us something here. A first application is the following.

Suppose $M_{\lambda 0}=\delta_{\lambda,0}$ — i.e. the 0th column of M is all zeros except for $M_{00}=1$. Then Perron-Frobenius implies (with a little work) that M will be a permutation matrix — i.e. there is some permutation π of P_+^k such that $M_{\lambda\mu}=\delta_{\mu,\pi\lambda}$, and $S_{\pi\lambda,\pi\mu}=S_{\lambda\mu}$. This nice fact applies directly to the $\mathcal{A}_\star$ and $\mathcal{D}_{odd}$ physical invariants of $A_1^{(1)}$.

This is proved by studying the powers $(M^t\,M)^L$ as L goes to infinity: its diagonal entries will grow exponentially with L, unless there is at most one nonzero entry on each row of M, and that entry equals 1.

More careful reasoning along those lines tells us about the other generic situation here. Namely, suppose $M_{\lambda 0}\neq 0$ only for $\lambda=0$ and $\lambda=J0$, and similarly for $M_{0\lambda}$ — i.e. the 0th row and column of M are all zeros except for $M_{J^i0,J^j0}=1$. Then the λth row (or column) of M will be identically 0 iff λ_1 is odd. Moreover, let λ,μ be any non-fixed-points of J, and suppose $M_{\lambda\mu}\neq 0$. Then

$$M_{\lambda\nu}=\begin{cases}1 & \text{if } \nu=\mu \text{ or } \nu=J\mu\\ 0 & \text{otherwise}\end{cases}$$

with a similar formula for $M_{\nu\mu}$. This applies to the $\mathcal{D}_{even}$ and $\mathcal{E}_7$ invariants of $A_1^{(1)}$.

Our final ingredient is the Galois symmetry (1.8.1) obeyed by S. Choose any Galois automorphism σ. It will correspond to some integer ℓ coprime to $2(k+2)$. From (1.8.1) and $M=SMS^*$ we get, for all λ,μ, the important relation

$$M_{\lambda\mu}=\epsilon_\sigma(\lambda)\,\epsilon_\sigma(\mu)\,M_{\lambda^\sigma,\mu^\sigma}\ . \tag{1.9.3}$$

From (1.9.3) and the positivity of M, we obtain the powerful *Galois selection rule*

$$M_{\lambda\mu}\neq 0 \quad\Longrightarrow\quad \epsilon_\sigma(\lambda)=\epsilon_\sigma(\mu)\ . \tag{1.9.4}$$

Next let us quickly sketch how these tools are used to obtain the $A_1^{(1)}$ classification. For details the reader should consult [29].

The first step will be to find all possible values of λ such that $M_{0\lambda}\neq 0$ or $M_{\lambda 0}\neq 0$. These λ are severely constrained. We know two generic possibilities: $\lambda_1=0$ (good for all k), and $\lambda=J0$ (good when $\frac{k}{2}$ is even). We now ask the question, what other possibilities for λ are there? Our goal is to prove (1.9.7). Assume $\lambda\neq 0,J0$, and write $a=\lambda_1+1$ and $n=k+2$.

There are only two constraints on λ which we will need. One is (1.9.1):

$$(a-1)\,(a+1)\equiv 0 \ (\mathrm{mod}\ 4n)\ . \tag{1.9.5}$$

More useful is the Galois selection rule (1.9.4), which we can write as $\sin(\pi\ell\frac{a}{n})\sin(\pi\ell\frac{1}{n})>0$, for all those ℓ. But a product of sines can be rewritten as a difference of cosines, so

$$\cos(\pi\,\ell\,\frac{a-1}{n})>\cos(\pi\,\ell\,\frac{a+1}{n})\ . \tag{1.9.6}$$

(1.9.6) is strong and easy to solve; the reader should try to find her own argument.

What we get is that, provided $n \neq 12, 30$, M obeys the strong condition

$$M_{\lambda 0} \neq 0 \text{ or } M_{0\lambda} \neq 0 \qquad \Longrightarrow \qquad \lambda \in \{0, J0\}\ . \tag{1.9.7}$$

Consider first **case 1**: $M_{\lambda 0} = \delta_{\lambda,0}$. From above, we know $M_{\lambda\mu} = \delta_{\mu,\pi\lambda}$ for some permutation π of P_+^k obeying $S_{\lambda\mu} = S_{\pi\lambda,\pi\mu}$. We know $\pi 0 = 0$; put $\mu := \pi(k-1,1)$. Then $\sin(\pi\frac{2}{n}) = \sin(\pi\frac{\mu_1+1}{n})$, and so we get either $\mu = (k-1,1)$ or $\mu = J(k-1,1)$. By T-invariance (1.9.1), the second possibility can only occur if $4 \equiv (n-2)^2 \pmod{4n}$, i.e. 4 divides n. But for those n, $\mathcal{D}_{\frac{n}{2}+1}$ is also a permutation matrix, so replacing M if necessary with the matrix product $M\,\mathcal{D}_{\frac{n}{2}+1}$, we can always require $\mu = (k-1,1)$, i.e. π also fixes $(k-1,1)$. It is now easy to show π must fix any λ, i.e. that M is the identity matrix $\mathcal{A}_{n-1}$.

The other possibility, **case 2**, is that both $M_{0,J0} \neq 0$ and $M_{J0,0} \neq 0$. (1.9.1) says $1 \equiv (n-1)^2 \pmod{4n}$, i.e. $\frac{n}{2}$ is odd. The argument here is similar to that of **case 1**, but with $(k-2,2)$ playing the role of $(k-1,1)$. We can show that $M_{(k-2,2),(k-2,2)} \neq 0$, except possibly for $k = 16$, where we find the exceptional $\mathcal{E}_7$. Otherwise we get $M = \mathcal{D}_{\frac{n}{2}+1}$.

For more general $X_\ell^{(1)}$ level k, the approach is

(i) to look at all the constraints on the $\lambda \in P_+^k$ for which $M_{0\lambda} \neq 0$ or $M_{\lambda 0} \neq 0$. Most important here are $TM = MT$ (which will always be some sort of norm selection rule) and the Galois selection rule (1.9.4). Generically, what we will find is that such a λ must equal $J0$ for some simple-current J, as in (1.9.7) for $A_1^{(1)}$.

(ii) Solve this generic case (in the $A_1^{(1)}$ classification, these were the physical invariants $\mathcal{A}_\star$, $\mathcal{D}_\star$ and $\mathcal{E}_7$).

(iii) Solve the nongeneric case. The worst of these are the orthogonal algebras at $k = 2$, as well as the places where conformal embeddings (see §1.7) occur.

(ii) has recently been completed for all simple X_ℓ, as has the $k = 2$ part of (iii). (i) is the main remaining task in the physical invariant classification for simple X_ℓ.

A natural question to ask is whether A-D-E has been observed in e.g. the $A_2^{(1)}$ classification. The answer is no, although the fusion graph theory of Di Francesco-Petkova-Zuber [62] is an attempt to assign to these physical invariants graphs reminiscent of the A-D-E Dynkin diagrams. Also, there is related work trying to understand the $A_2^{(1)}$ classification in terms of subgroups of $\mathrm{SU}_3(\mathbb{C})$ (as opposed to $\mathrm{SU}_2(\mathbb{C})$ for $A_1^{(1)}$) — see e.g. [35]. Finding the $A_3^{(1)}$, $A_4^{(1)}$,... classifications would permit the clarification and testing of this vaguely conjectured relation between the $A_n^{(1)}$ physical invariants, and singularities $\mathbb{C}^{n+1}/G$ for G a finite subgroup of $\mathrm{SU}_{n+1}(\mathbb{C})$.

However, a few years ago Philippe Ruelle was walking in a library in Dublin. He spotted a yellow book in the math section, called *Complex Multiplication* by Lang. A strange title for a book by Lang! After all, there can't be all that much even Lang could really say about complex multiplication! Ruelle flipped it to a random page, which turned out to be p.26. On there he found what we would call the Galois selection rule for $A_2^{(1)}$, analysed and solved for the cases where $k+3$ is coprime to 6. Lang however didn't know about physical invariants; he was reporting on work by Koblitz and Rohrlich on decomposing the Jacobians of the Fermat curve $x^n + y^n = z^n$ into their prime pieces, called 'simple factors' in algebraic geometry. n here corresponds to $k + 3$. Similarly, Itzykson discovered traces of the $A_2^{(1)}$ exceptionals — these occur when $k + 3 = 8, 12, 24$ — in the Jacobian of $x^{24} + y^{24} = z^{24}$. See [2] for further observations along these lines. These 'coincidences' are still far from understood. Nor is it known if, more generally, the $A_\ell^{(1)}$ level k classification will somehow be related to the hypersurface $x_1^n + \cdots + x_\ell^n = z^n$, for $n = k + \ell + 1$.

The $(u(1) \oplus \cdots \oplus u(1))^{(1)}$ classification has connections to rational points on Grassmannians. The Grassmannian is (essentially) the moduli space for the Narain compactifications of the (classical) lattice string. It would be very interesting to interpret other large families of physical invariants as special points on other moduli spaces.

These new connections relating various physical invariant classifications to other areas of math seem to indicate that although the physical invariant classifications are difficult, they could be well worth the effort and be of interest outside RCFT. Once the physical invariant lists are obtained, we will still have the fascinating task of explaining and developing all these mysterious connections. These thoughts keep me going!

Another motivation for completing these lists comes from their relation to subfactor theory in von Neumann algebras[18]. These algebras (see e.g. [22]) can be thought of as symmetries of a (generally infinite) group. Their building blocks are called *factors.* Jones initiated the combinatorial study of *subfactors* N of M (i.e. inclusions $N \subseteq M$ where M, N are factors), relating it to e.g. knots, and for this won a Fields medal in 1990. Jones assigned to each subfactor $N \subseteq M$ a numerical invariant called an 'index', a sort of (generally irrational) ratio of dimensions. Graphs (called principal and dual principal) are also associated to subfactors. A much more refined subfactor invariant, called a 'paragroup', has been introduced by Ocneanu. It is essentially equivalent to a (2+1)-dimensional topological field theory. Moreover, any RCFT can be assigned a paragroup, and any paragroup (via a process called asymptotic inclusion which is akin to Drinfeld's quantum doubling of Hopf algebras) yields an RCFT. See [22] for details.

Böckenhauer-Evans [4] have recently developed this much further, and have clarified the fusion graph ↔ physical invariant relation. The fusion graphs will correspond to subfactor principal graphs. In the work of Di Francesco-Petkova-Zuber, that relation seems to be only empirical (i.e. nonconceptual).

Subfactor theory together with singularity theory is our best hope at present for understanding and generalising the A-D-E meta-pattern.

---

[18] For reasons of necessity, in the following discussion I'll take more liberties than usual in the presentation.

Part 2. Monstrous Moonshine

2.1. INTRODUCTION

In 1978, John McKay made a very curious observation. One of the well-known[19] functions of classical number theory is the j-function[20], given by

$$j(\tau) := \frac{(1+240\sum_{n=1}^{\infty}\sigma_3(n)\,q^n)^3}{q\prod_{n=1}^{\infty}(1-q^n)^{24}} = \frac{\Theta_{E_8}(\tau)^3}{\eta(\tau)^{24}}$$
$$= q^{-1} + 744 + 196\,884\,q + 21\,493\,760\,q^2 + 864\,299\,970\,q^3 + \cdots \tag{2.1.1}$$

Here as elsewhere in this paper, $q = \exp[2\pi\mathrm{i}\,\tau]$. Also, $\sigma_3(n) = \sum_{d|n} d^3$, Θ_{E_8} is the theta function of the E_8 root lattice, and η is the Dedekind eta. What is important here are the values of the first few coefficients. What McKay noticed was that $196\,884 \approx 196\,883$. Closer inspection shows $21\,493\,760 \approx 21\,296\,876$, and $864\,299\,970 \approx 842\,609\,326$. In fact,

$$196\,884 = 196\,883 + 1 \tag{2.1.2a}$$
$$21\,493\,760 = 21\,296\,876 + 196\,883 + 1 \tag{2.1.2b}$$
$$864\,299\,970 = 842\,609\,326 + 21\,296\,876 + 2\cdot 196\,883 + 2\cdot 1 \tag{2.1.2c}$$

The numbers on the right-side are the dimensions of the smallest irreducible representations of the Monster finite simple group $\mathbb{M}$ (in 1978 it still wasn't certain that $\mathbb{M}$ even existed so back then these numbers were merely conjectural). The same game could be played with other coefficients of the j-function. With numbers so large, it seemed to him doubtful that this numerology was merely a coincidence. On the other hand, it was hard to imagine any deep conceptual connection between the Monster and the j-function: they seem completely unrelated.

In November 1978 he mailed the 'McKay equation' (2.1.2a) to John Thompson. At first Thompson dismissed this as nonsense, but after checking the next few coefficients he became convinced. He then added a vital piece to the puzzle. It should be well-known that when one sees a nonnegative integer, it often helps to try to interpret it as the dimension of some vector space. Essentially, that is what McKay was proposing here. (2.1.2) are really hinting that there is a 'graded' representation V of $\mathbb{M}$:

$$V = V_{-1} \oplus V_1 \oplus V_2 \oplus V_3 \oplus \cdots$$

where $V_{-1} = \rho_0$, $V_1 = \rho_1 \oplus \rho_0$, $V_2 = \rho_2 \oplus \rho_1 \oplus \rho_0$, $V_3 = \rho_3 \oplus \rho_2 \oplus \rho_1 \oplus \rho_1 \oplus \rho_0 \oplus \rho_0$, etc, where ρ_i are the irreducible representations of $\mathbb{M}$ (ordered by dimension), and that

$$j(\tau) - 744 = \dim_q(V) := \dim(V_{-1})\,q^{-1} + \sum_{i=1}^{\infty}\dim(V_i)\,q^i\ , \tag{2.1.3}$$

---

[19] 'Well-known' is math euphemism for 'a basic result of which until recently we were utterly ignorant.' As Conway later said, "the j-function was 'well-known' to other people, but not 'well-known' to me."

[20] This and other technical terms used in this introduction will be carefully explained in the following subsections. This section is merely offered as a quick overview.

the graded dimension of V.

Thompson suggested that we twist $\dim_q(V)$, i.e. that more generally we consider the series (now called the *McKay-Thompson series*)

$$T_g(\tau) := \mathrm{ch}_{V,q}(g) = \mathrm{ch}_{V_{-1}}(g)\, q^{-1} + \sum_{i=1}^{\infty} \mathrm{ch}_{V_i}(g)\, q^i \ , \tag{2.1.4}$$

for each element $g \in \mathbb{M}$. The point is that, for any group representation ρ, the character value $\mathrm{ch}_\rho(id.)$ equals the dimension of ρ, and so $T_{id.}(\tau) = j(\tau) - 744$ and we recover (2.1.2) as special cases. But there are many other possible choices of $g \in \mathbb{M}$. Thompson couldn't guess what these functions T_g would be, but he suggested that they too might be interesting. This is a nice thought: when we see a positive integer, we should try to interpret it as a dimension of a vector space; if there is a symmetry present, then it may act on the vector space — i.e. our vector space may carry a representation of that symmetry group — in which case we can apply the Thompson trick and see what if any significance the other character values have in our context.

Conway and Norton [12] did precisely what Thompson asked. Conway called it "one of the most exciting moments in my life" [11] when he opened Jacobi's foundational (but 150 year old!) book on elliptic and modular functions and found that the first few terms of the McKay-Thompson series agreed perfectly with the first few terms of certain special functions, namely the Hauptmoduls of various genus 0 modular groups. Monstrous Moonshine was officially born.

The word 'moonshine' here is English slang for 'unsubstantial or unreal'. It was chosen by Conway to convey as well the feeling that things here are dimly lit, and that Conway-Norton were 'distilling information illegally' from the Monster character table.

In fact the first incarnation of Moonshine goes back to Andrew Ogg in 1975. He was in France describing his result that the primes p for which the group $\Gamma_0(p)+$ has genus 0, are $\{2, 3, 5, 11, 13, 17, 19, 23, 29, 31, 41, 47, 59, 71\}$. $\Gamma_0(p)+$ is the group generated by $\begin{pmatrix} 0 & 1 \\ -p & 0 \end{pmatrix}$ and $\Gamma_0(p)$, and is the normaliser of $\Gamma_0(p)$ in $\mathrm{SL}_2(\mathbb{R})$ (this sentence will make a little more sense after §2.3, but it isn't important here to understand it). He also attended a lecture by Jacques Tits, who was describing a newly conjectured simple group. When Tits wrote down the prime decomposition of the order of that group (see (2.2.1) below), Ogg noticed its prime factors precisely equalled his list of primes. Presumably as a joke, he offered a bottle of Jack Daniels' whisky to the first person to explain the coincidence.

The next step was accomplished by Griess in 1980, with the construction of the Monster[21] $\mathbb{M}$, and with it the proof that the conjectured character table for $\mathbb{M}$ was correct. Griess did this by explicitly constructing the 196883-dimensional representation ρ_1; it turns out to have a (commutative nonassociative) algebra structure, now called the *Griess algebra*. Though this paper was clearly important, the construction was artificial and 100 pages long: since the Monster is presumably a natural mathematical object (see §2.2), an elegant construction for it should exist. This was ultimately accomplished in the mid 1980s with the construction by Frenkel-Lepowsky-Meurman [23] of the Moonshine module

---

[21] Griess also came up with the symbol for the Monster; Conway came up with the name.

$V^\natural$ and its interpretation by Borcherds as a *vertex operator algebra*. The Griess algebra appears naturally in $V^\natural$, as we shall see. $V^\natural$ does indeed seem to be a 'natural' mathematical structure, and $\mathbb{M}$ is its automorphism group: in fact $V^\natural$ is the graded representation V of $\mathbb{M}$ conjectured by McKay and Thompson.

Connections with physics (CFT) go back to Dixon-Ginsparg-Harvey [19] in 1988, in a paper titled "Beauty and the beast: Superconformal symmetry in a Monster module". The Moonshine module $V^\natural$ can be interpreted as the string theory for a $\mathbb{Z}_2$-orbifold of free bosons compactified on the torus $\mathbb{R}^{24}/\Lambda_{24}$ (Λ_{24} is the Leech lattice). Many aspects of Moonshine make complete sense within CFT, but some (e.g. the genus zero property) remain more obscure. (Though in 1987 Moore speculated that the 0-genus of $\Gamma_0(a)+$ could be related to the vanishing of the cosmological constant in certain string theories related to $\mathbb{M}$, and Tuite [57] related genus-zero with the conjectured uniqueness of $V^\natural$.) Nevertheless this helps make the words of Dyson ring prophetic: "I have a sneaking hope, a hope unsupported by any facts or any evidence, that sometime in the twenty-first century physicists will stumble upon the Monster group, built in some unsuspected way into the structure of the universe" [21].

Finally, in 1992 Borcherds [5] completed the proof of the Conway-Norton conjectures by showing $V^\natural$ is the desired representation V. The full conceptual relationship between the Monster and the Hauptmoduls (like j) seems to remain 'dimly lit', although much progress has been realised. This is a subject where it is much easier to conjecture than to prove, and we are still awash in unresolved conjectures.

McKay also noticed in 1978 that similar coincidences hold if $\mathbb{M}$ and $j(\tau)$ are replaced with the Lie group $E_8(\mathbb{C})$ and $(qj(q))^{\frac{1}{3}} = 1 + 248q + \cdots$. This turns out to be much easier to explain, and in 1980 both Kac and Lepowsky remarked that the unique level 1 highest-weight representation of the affine algebra $E_8^{(1)}$ has graded dimension $(qj(q))^{\frac{1}{3}}$.

Moonshiners have a little chip on their shoulders. Modern math, they say, tends to be a little too infatuated with the pursuit of generalisations for generalisations' sake. Surely a noble goal for math is to find interesting and fundamentally new theorems. It can be argued that both history and common-sense suggest that to this end it is most profitable to look simultaneously at both exceptional structures and generic structures, to understand the special features of the former in the context of the latter, and to be led in this way to a new generation of exceptional and generic structures. Moonshiners would sympathise with those biologists who study the duck-billed platypus and lungfish rather than hide them in the closet as monsters: BECAUSE they appear to be unique, those animals presumably have much to teach us about our general understanding of evolution, etc.

It often seems to people that Moonshine can't be very deep: the Conway-Norton conjectures seem to be so finite[22] and specialised. There only are 171 distinct McKay-Thompson series T_g in Monstrous Moonshine, after all. The whole point though is to try to understand *why* the Monster and the Hauptmoduls are so related, and then to try to extend and apply this understanding to other contexts. Moonshine is still young, and our understanding remains incomplete. But already math has benefitted: e.g. we now have a natural definition of $\mathbb{M}$ (as the automorphism group of $V^\natural$), and Moonshine helped lead us to the rich structures of generalised Kac-Moody algebras and vertex operator algebras.

---

[22] Indeed the Moonshine conjectures *are* finite (it is enough to check the first 1200 coefficients), and a slightly weaker form was quickly proved on a computer by Atkin, Fong and Smith [54]. However this sort of argument adds no light to Moonshine, and tells us nothing of V except that it exists.

We will see that Moonshine involves the interplay between *exceptional structures* such as the number 24, the Leech lattice Λ_{24}, the Monster group $\mathbb{M}$, and the Moonshine module $V^\natural$, and *generic structures* such as modular functions, vertex operator algebras, generalised Kac-Moody algebras, and conformal field theories. The following sections will introduce the reader to many of these structures, as we use Moonshine as another happy excuse to take a second little tour through modern mathematics.

2.2. INGREDIENT #1: FINITE SIMPLE GROUPS AND THE MONSTER

A readable introduction to the basics of finite group representation theory is [25]. The finite simple groups are described in [33]; see also [11]. Group representations were introduced in §1.3.

A *normal* subgroup H of a group is one obeying $gHg^{-1} = H$ for all $g \in G$. These are important because the set G/H of 'cosets' gH has a natural group structure precisely when H is normal. Every group has two trivial normal subgroups: itself and $\{1\}$. If these are the only normal subgroups, the group is called *simple*. It is conventional to regard the trivial group $\{1\}$ as not simple (just as 1 is conventionally regarded as not prime). An alternate definition of a (finite) simple group G is that if $\varphi : G \to H$ is any group homomorphism (i.e. structure-preserving map: $\varphi(gg') = \varphi(g)\varphi(g')$), then φ is either constant (i.e. $\varphi(G) = \{1\}$), or φ is one-to-one.

The importance of simple groups is provided by the *Jordan-Hölder Theorem*. By a 'composition series' for a group G, we mean a nested sequence

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_k \supset H_{k+1} = \{1\}$$

of groups such that H_i is normal in H_{i-1}, and H_{i-1}/H_i (called a 'composition factor') is simple. Any finite group G has at least one composition series. If $H'_0 \supset \cdots \supset H'_{\ell+1} = \{1\}$ is a second composition series for G, then Jordan-Hölder says that $k = \ell$ and, up to a reordering π, the simple groups H_{i-1}/H_i and $H'_{\pi j-1}/H'_{\pi j}$ are isomorphic.

For example, the cyclic group $\mathbb{Z}_n$ of order(=size) n — you can think of it as the integers modulo n under addition — is simple iff n is prime. Consider the group $\mathbb{Z}_{12} = \langle 1 \rangle$. Two composition series are

$$\mathbb{Z}_{12} \supset \langle 2 \rangle \supset \langle 4 \rangle \supset \langle 0 \rangle$$
$$\mathbb{Z}_{12} \supset \langle 3 \rangle \supset \langle 6 \rangle \supset \langle 0 \rangle$$

corresponding to composition factors $\mathbb{Z}_2$, $\mathbb{Z}_2$, $\mathbb{Z}_3$, and $\mathbb{Z}_3$, $\mathbb{Z}_2$, $\mathbb{Z}_2$. Of course this is consistent with Jordan-Hölder. This is reminiscent of the fact that $2 \cdot 2 \cdot 3 = 3 \cdot 2 \cdot 2$ are both prime factorisations of 12.

There is some value to regarding finite groups as a massive generalisation of the notion of number. The number n can be identified with the cyclic group $\mathbb{Z}_n$. The divisor of a number corresponds to a normal subgroup, so a prime number corresponds to a simple group. The Jordan-Hölder Theorem generalises the uniqueness of prime factorisations. That you can build up any number by multiplying primes, is generalised to building up a group by semi-direct products (more generally, by group extensions): if H is a normal subgroup of G, then G will be an extension of H by the quotient group G/H.

Note however that $\mathbb{Z}_6 \times \mathbb{Z}_2$ and $\mathfrak{S}_3 \times \mathbb{Z}_2$ — both different from $\mathbb{Z}_{12}$ — will also have $\mathbb{Z}_2, \mathbb{Z}_2, \mathbb{Z}_3$ as composition factors: unlike for numbers, 'multiplication' here does not give a unique answer. The semidirect product $\mathbb{Z}_2 \rtimes \mathbb{Z}_2$ can equal either $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$, depending on how the product is taken. More precisely, the notation $G \rtimes G'$ means a group where every element can be written uniquely as a pair (g, g'), for $g \in G$ and $g' \in G'$, and where the group operation is $(g, g')(h, h') = (\text{stuff}, gh)$.

Thus simple groups have an importance for group theory approximating what primes have for number theory. One of the greatest accomplishments of twentieth century math is surely the classification of the finite simple groups. (On the other hand, group extensions turn out to be technically quite difficult and leads one into group cohomology.) This work, completed in the early 1980s (although gaps are continually being discovered and filled in the arguments), runs to approximately 15 000 journal pages, spread over 500 individual papers, and is the work of a whole generation of group theorists. A modern revision is currently underway (see e.g. [34]) to simplify the proof and find and fill all gaps, but the final proof is still expected to be around 4000 pages long. The resulting list is:

- the cyclic groups $\mathbb{Z}_p$ (p a prime);
- the alternating groups $\mathfrak{A}_n$ for $n \geq 5$;
- 16 families of Lie type;
- 26 sporadics.

We've already met the cyclic groups. The alternating group $\mathfrak{A}_n$ consists of the even permutations in the symmetric group $\mathfrak{S}_n$, and so has order(=size) $\frac{1}{2}\,n!$. The groups of Lie type are essentially Lie groups defined over finite fields[23] $\mathbb{F}_q$ (such as $\mathbb{Z}_p$), sometimes 'twisted' in certain senses. The simplest example is $\mathrm{PSL}_n(\mathbb{F}_q)$, which consists of the $n \times n$ matrices with entries in $\mathbb{F}_q$, with determinant 1, quotiented out by the centre of $\mathrm{SL}_n(\mathbb{F}_q)$ (namely the scalar matrices $\mathrm{diag}(a, a, \ldots, a)$ for $a^n = 1$) (except for $\mathrm{PSL}_2(\mathbb{Z}_2)$ and $\mathrm{PSL}_2(\mathbb{Z}_3)$, which aren't simple).

Note that the determinant $|\rho(g)|$ for any representation ρ of any (noncyclic) simple group must be 1, otherwise we would violate the homomorphism definition of simple group (try to see why). Also, the centre of any (noncyclic) simple group must be trivial (why?). The smallest noncyclic simple group is $\mathfrak{A}_5$, with order 60.[24] It is the same as (isomorphic to) $\mathrm{PSL}_2(\mathbb{Z}_5)$ and $\mathrm{PSL}_2(\mathbb{F}_4)$, and can also be expressed as the group of all rotations (reflections have determinant -1 and so cannot belong to any simple group) of $\mathbb{R}^3$ that bring a regular icosahedron back to itself.

---

[23] There is a finite field with q elements, iff q is a power of a prime. For each such q, there is only 1 field of that size. The field with prime p elements is the integers taken mod p.

[24] This implies, incidentally, that if G and H are any two groups with the same order below 60, then they will have the same composition factors.

The smallest sporadic group is the Mathieu group M_{11}, order 7920, discovered in 1861[25]. The largest is the Monster $\mathbb{M}$, conjectured by Fischer and Griess in 1973 and finally proved to exist by Griess in 1980. Its order is[26]

$$\|\mathbb{M}\| = 2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \times 10^{53} \ . \quad (2.2.1)$$

20 of the 26 sporadics are involved in (i.e. are quotients of subgroups of) the Monster. Some relations among $\mathbb{M}$, the Leech lattice Λ_{24} and the largest Mathieu group M_{24} are given in Chapters 10 and 29 of [13].

Moonshine hints at a tantalising connection between the classification of finite simple groups, and the classification of RCFTs discussed in Part 1. Speculates [23] (page xli): "One can certainly hope for a uniform description of the finite simple groups as automorphism groups of certain vertex operator algebras — or conformal quantum field theories. If such a quantum field theory could somehow be attached a priori to a finite simple group, the classification of such theories, a problem of great current interest among string theorists, might some day be part of a new approach to the classification of the finite simple groups. On the other hand, can the known classification of the finite simple groups help in the classification of conformal field theories?"

2.3. INGREDIENT #2: MODULAR FUNCTIONS AND HAUPTMODULS

A readable introduction to some of the topics discussed in this section is [16,42,60].

We know from complex analysis that the group $\mathrm{SL}_2(\mathbb{R})$ of 2×2 matrices with real entries and determinant 1, acts on the upper-half plane $\mathcal{H} = \{\tau \in \mathbb{C} \,|\, \mathrm{Im}(\tau) > 0\}$ by fractional linear (or *Möbius*) transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau := \frac{a\tau + b}{c\tau + d} \ . \qquad (2.3.1)$$

For example the matrix $S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponds to the function $\tau \mapsto -1/\tau$, while the matrix $T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ corresponds to the translation $\tau \mapsto \tau + 1$. Since $\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ correspond to the same Möbius transformation, strictly speaking our group here is $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm I\}$.

The only reason this action (2.3.1) of the 2×2 matrices on complex numbers (or more precisely the Riemann sphere $\mathbb{C} \cup \{\infty\}$) might not look strange to us, is because familiarity breeds numbness. What we really have is a natural action of $n \times n$ matrices on $\mathbb{C}^n$, and this induces their action on $\mathbb{C}^{n-1}$ (together with a codimension-2 set of 'points at infinity') by interpreting $\mathbb{C}^n$ as projective coordinates for $\mathbb{C}^{n-1}$. Specialising to $n = 2$ gives us (2.3.1). In projective geometry, 'parallel lines' intersect at ∞. Projective coordinates allow one to treat 'finite' and 'infinite' points on an equal footing.

---

[25] Although his arguments apparently weren't very convincing. In fact some people, including the Jordan of Jordan-Hölder fame, argued in later papers that the largest of Mathieu's sporadic groups couldn't exist.

[26] The inquisitive reader, hungry for more 'coincidences', may have noticed that 196883 and 21296876 — see (2.1.2) — exactly divide the order of the Monster. Indeed this will hold for any finite group: the dimensions of the irreducible representations of a finite group will always divide its order.

Consider $\Gamma := \mathrm{SL}_2(\mathbb{Z})$, the subgroup of $\mathrm{SL}_2(\mathbb{R})$ consisting of the matrices with integer entries. It can be shown that it is generated by S and T (in other words, every matrix $\alpha \in \Gamma$ can be expressed as a monomial in S and T). For reasons that will be clear shortly, consider the extended upper-half plane $\overline{\mathcal{H}} := \mathcal{H} \cup \{\mathrm{i}\infty\} \cup \mathbb{Q}$ — the extra points $\{\mathrm{i}\infty\} \cup \mathbb{Q}$ are called *cusps*. Γ acts on $\overline{\mathcal{H}}$ (e.g. S interchanges 0 and $\mathrm{i}\infty$). By a *modular function* for Γ, we mean a meromorphic function $f : \overline{\mathcal{H}} \to \mathbb{C}$, symmetric with respect to Γ: i.e. $f(\alpha(\tau)) = f(\tau)$ for all $\alpha \in \Gamma$. Note that we require f to be meromorphic at the cusps (e.g. polynomials are meromorphic at $\mathrm{i}\infty$, but e^z is not).

It is not obvious why modular functions should be interesting, but in fact they are one of the most fundamental notions in modern number theory (see the last paragraph of §1.6). For example, consider the question of writing numbers as sums of squares. We can write $5 = 1^2 + (-2)^2 = (-1)^2 + 1^2 + 0^2 + 1^2 + (-1)^2$, to give a couple of trivial examples. Let $N_n(k)$ be the number of ways we can write the integer n as a sum of k squares, counting order and signs. For example $N_5(1) = 0$ (since 5 is not a perfect square), $N_5(2) = 8$ (since $5 = (\pm 1)^2 + (\pm 2)^2 = (\pm 2)^2 + (\pm 1)^2$), $N_5(3) = 24$, etc. Their generating functions are:[27]

$$\sum_{n=0}^{\infty} N_n(k)\, q^n = (\theta_3(q))^k \ ,$$

where

$$\theta_3(q) = 1 + 2q + 2q^4 + \cdots = \sum_{n \in \mathbb{Z}} q^{n^2}$$

is called a *theta function*. It turns out that θ_3 transforms nicely with respect to Γ, once we make the change-of-variables $q = \exp[\pi \mathrm{i} \tau]$. This takes work to show. For example, θ_3 is clearly invariant under the action of $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, and a little work (from e.g. Poisson summation) shows that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ takes $\theta_3(\tau)$ to $\sqrt{\frac{\tau}{\mathrm{i}}}\,\theta_3(\tau)$. θ_3 is not precisely a modular function (it is a 'modular form of weight $\frac{1}{2}$' for $\Gamma_0(4)$), but this simple example illustrates the point that Γ (and related groups) appear throughout number theory. More on this shortly.

That important change-of-variables $q = \exp[\pi \mathrm{i} \tau]$ was introduced by Jacobi early last century, in his analysis of 'elliptic integrals'. The theory is beautiful and poorly remembered today, which is very disappointing considering how much of modern math was touched by it. I strongly recommend the book [10], written over a century ago; the style and motivation of math in our century is different from that in Jacobi's, and we've lost a little in motivation what we've gained in power. I'll briefly sketch Jacobi's theory.

---

[27] A fundamental principle in math is: whenever you have a subscript with an infinite range, make a power series (called a *generating function*) out of it.

Just as we could develop a theory of 'circular functions' (i.e. sine etc.) starting from the integral $s(a) = \int_0^a \frac{dx}{\sqrt{1-x^2}}$, so can we develop a theory of 'elliptic functions' starting from the 'elliptic integral' $F(k,a) = \int_0^a \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$. Inverting $s(a)$ gives a function both more useful and with nicer properties than $s(a)$: we call it $\sin(u)$. Similarly, for any k the elliptic function $\mathrm{sn}(k,u)$ is defined by $u = F(k, \mathrm{sn}(k,u))$. Just as we can define a numerical constant π by $\sin(\frac{1}{2}\pi) = 1$ (i.e. $\frac{1}{2}\pi = \int_0^1 \frac{dx}{\sqrt{1-x^2}}$), we get a function $K(k) = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$. Just as $\sin(u)$ has period $4(\frac{1}{2}\pi)$, so has sn u-period $4K(k)$. sn also turns out to have u-period $4\mathrm{i}\,K(k')$ where $k' = \sqrt{1-k^2}$ — today we take this as the starting point and define an elliptic function to be doubly periodic (see [42] or Cohen in [60]).

The theta functions aren't elliptic functions, but they are closely related, and e.g. sn can be written as a quotient of them. In Jacobi's language, we have

$$\theta_3\left(\frac{\mathrm{i}\,K(k')}{K(k)}\right) = \sqrt{\frac{2K(k)}{\pi}}\ .$$

The 'modular transformation' $\tau \mapsto \frac{-1}{\tau}$ corresponds to interchanging the 'modulus' k with the 'complementary modulus' k', and thus is completely natural in Jacobi's theory. The important formula $\theta_3(\frac{-1}{\tau}) = \sqrt{\frac{\tau}{\mathrm{i}}}\,\theta_3(\tau)$ is trivial here.

A certain interpretation of modular functions also indicates their usefulness, and played an important role in Part 1. A *torus* is something that looks like the surface of a bagel, at least as far as its topology is concerned. For example, the Cartesian product $S^1 \times S^1$ of circles is a torus (think of one circle being the contact-circle of the bagel with the table on which it rests, then from each point on that horizontal circle imagine placing a vertical circle perpendicular to it, like a rib; together all these ribs fill out the bagel's surface). A more sophisticated example of a torus is an elliptic curve (a complex curve of the form $y^2 = ax^3 + bx^2 + cx + d$ and a special point on it playing the role of 0). A final example is the quotient $\mathbb{C}/\Lambda$ of the complex plane $\mathbb{C}$ with a 2-dimensional lattice Λ (we saw lattices in §1.6; Λ here will be a discrete doubly-periodic set of points in $\mathbb{C}$, containing 0). It turns out that certain equivalence classes of tori (e.g. with respect to conformal or complex-analytic equivalence) always contain a representative torus of the form $\mathbb{C}/\Lambda$, where Λ consists of all points $\mathbb{Z} + \mathbb{Z}\,\tau$, for some $\tau \in \mathcal{H}$. (Incidentally, the cusps correspond to degenerate tori.) In other words, these equivalence classes are parametrized by complex numbers τ in $\mathcal{H}$. So if we have a complex-valued function F on the set of all tori, which is e.g. conformally invariant (an example is the genus-one partition function $\mathcal{Z}$ in conformal field theories — see §1.1), then we can consider F as a well-defined function $F : \mathcal{H} \to \mathbb{C}$. However, it turns out that different points τ in $\mathcal{H}$ correspond to the same equivalence class of tori: e.g. the lattice for τ is the same as that for $\tau + 1$, and these are a rescaling of that for $-1/\tau$. Thus $F(\tau) = F(\tau + 1) = F(-1/\tau)$, because $\tau, \tau + 1, -1/\tau$ all represent equivalent tori. Since $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$ generate $\mathrm{PSL}_2(\mathbb{Z})$, what in fact we find is that F has Γ as its group of symmetries. One often says that Γ is the 'modular group of the torus', and that the orbit space $\Gamma\backslash\mathcal{H}$ is the 'moduli space' of (conformal equivalence classes of) tori. $\mathcal{H}$ is called its 'Teichmüller space' or 'universal cover'. This is exactly analogous to $S^1 = \mathbb{R}/\mathbb{Z}$: $\mathbb{R}$ is its universal cover and $\mathbb{Z}$ is its 'modular group' (or 'mapping class group'). Another example: the Teichmüller space for (conformal equivalence classes of) 'pair-of-pants', or equivalently a disc minus two open interior disks, is $\mathbb{R}^3$ (an ordered triple), while its modular group is the symmetric group $\mathfrak{S}_3$ and its moduli space consists of unordered triples. Incidentally, we write $\Gamma\backslash\mathcal{H}$ instead of $\mathcal{H}/\Gamma$ because the group Γ acts on $\mathcal{H}$ 'on the left'. A good introduction to the geometry here is [55].

In any case, a surprising number of innocent-looking questions in number theory can be dragged (usually with effort) into the richly developed realm of elliptic curves and modular functions, and it is there they are often solved. For instance, we all know the ancient Greeks were interested in Pythagorean triples: find all integer solutions a, b, c to $a^2 + b^2 = c^2$, i.e. find all integer (or if you prefer, rational) right-angle triangles. They solved this by elementary means: choose any integers (or rationals) x, y and put $u = \frac{x^2-y^2}{x^2+y^2}$, $v = \frac{2xy}{x^2+y^2}$; then $u^2 + v^2 = 1$ and (multiplying by the denominator) this gives all Pythagorean triples.

There are two ways of extending this problem. One is to ask which $n \in \mathbb{Z}$ can arise as areas of these rational right-angle triangles. It turns out $n = 5$ is the smallest one: $a = \frac{3}{2}$, $b = \frac{20}{3}$, $c = \frac{41}{6}$ works ($5 = \frac{1}{2}(\frac{3}{2})(\frac{20}{3})$ and $(\frac{3}{2})^2 + (\frac{20}{3})^2 = (\frac{41}{6})^2$). This is a hard problem — just try to show $n = 1$ cannot work. $n = 157$ turns out to work: the simplest triangle has a and b as quotients of integers of size around 10^{25}, and c as the quotient of integers around 10^{47}. Although this problem was studied by the ancient Greeks and also by the Arabs in the 10th century, it was finally cracked in the 1980s. It was solved by first translating it into the question of whether the elliptic curve $y^2 = x^3 - n^2x$ has infinitely-many rational points, and then applying all the rich 20th century machinery to answering that question.

The other continuation of the Pythagorean triples question is more famous: find all integer solutions to $a^n + b^n = c^n$ (or equivalently all rational solutions to $a^n + b^n = 1$). 350 years ago Fermat wrote in the margin of the book he was reading (the book was describing the Greek solution to Pythagorean triples) that he had found a "truly marvelous" proof that for $n > 2$ there are no nontrivial solutions, but that the margin was too narrow to contain it. This result came to be known as 'Fermat's Last Theorem'[28] and despite considerable effort no one has succeeded in rediscovering his proof. Most people today believe that Fermat soon realised his 'proof' wasn't valid, otherwise he would have alluded to it in later letters. In any case, a very long and complicated proof was finally achieved in the 1990s: the 'Taniyama conjecture' says that a certain function associated to any elliptic curve over $\mathbb{Q}$ will be modular; if $a^n + b^n = c^n$ for some $n > 2$, then the elliptic curve $y^2 = x^3 + (a^n - b^n)x^2 - a^nb^n$ will violate the Taniyama conjecture; finally Wiles proved the Taniyama conjecture is true.

To most mathematicians, the 'area-n problem' and 'Fermat's Last Theorem' are interesting only because they can be related to elliptic curves and modular forms — it's easy to ask hard questions in math, but most questions tend to be stale. Number theory is infatuated with modular stuff because (in increasing order of significance) (a) it's exceedingly rich, with lots of connections to other areas of math and math phys; (b) it's a battleground on which many innocent-looking but hard-to-crack problems can be slain; and (c) last generation's number theorists also worked on modular stuff.

---

[28] It was called his 'Last Theorem' because it was the last of his 48 margin notes to be proved by other mathematicians — another one is discussed in Section 1.8. The story of Fermat's Last Theorem is a fascinating one, but alas this footnote is too small to do it credit. See for instance the excellent book [53].

In any case, modular functions turn out to be important for math (and mathematical physics) even though they may at first glance look artificial. Poincaré explained how to study them. He said to look at the orbits of $\overline{\mathcal{H}}$ with respect to Γ. For example, one orbit, hence one point in $\Gamma\backslash\overline{\mathcal{H}}$, contains all cusps. We write this as $\Gamma\backslash\overline{\mathcal{H}}$, and give it the natural topological structure (i.e. 2 points $[\tau],[\tau'] \in \Gamma\backslash\overline{\mathcal{H}}$ are considered 'close' if the 2 sets $\Gamma\tau, \Gamma\tau'$ nearly overlap). Note first that by applying T repeatedly, every point in $\mathcal{H}$ corresponds to a point in the vertical strip $-\frac{1}{2} \leq \mathrm{Re}(\tau) \leq \frac{1}{2}$ — in fact to a unique point in that strip, if we avoid the two edges. S is an inversion through the unit circle, so it permits us to restrict to those points in the vertical strip which are distance at least 1 from the origin. The resulting region R is called a fundamental region for Γ. Apart from the boundary of R, every Γ-orbit will intersect R in one and only one point.

What should we do about the boundary? Well, the edge $\mathrm{Re}(\tau) = -\frac{1}{2}$ gets mapped by T to the edge $\mathrm{Re}(\tau) = \frac{1}{2}$, so we should identify (=glue together) these. The result is a cylinder running off to infinity, with a strange lip at the bottom. S tells us how we should close that lip: identify $\mathrm{i}e^{\mathrm{i}\theta}$ and $\mathrm{i}e^{-\mathrm{i}\theta}$. This seals the bottom of the cylinder, so we get an infinitely tall cup with a strangely puckered base. In fact the top of this cup is also capped off, by the cusp $\mathrm{i}\infty$. So what we have (topologically speaking) is a *sphere*. It does not look like a smooth sphere, but in fact it inherits the smoothness of $\mathcal{H}$.

Incidentally, topological manifolds of dimension ≤ 3 always have a unique compatible smooth structure. 'Topological structure' means you can speak of continuity or closeness, 'smooth structure' means you can also do calculus. On the other hand, $\mathbb{R}^4$ has infinitely many smooth structures compatible with its topological structure; mysteriously, all other Euclidean spaces $\mathbb{R}^n$ have a unique smooth structure! Thus both mathematics and physics single out 4-space. Coincidence???

So anyways, what this construction of $\Gamma\backslash\overline{\mathcal{H}}$ means is that a modular function can be reinterpreted as a meromorphic complex-valued function on this sphere. This is very useful, because our undergraduate complex variables class taught us all about meromorphic complex-valued functions f on the Riemann sphere $\mathbb{C} \cup \infty$. There are many meromorphic functions on $\mathbb{C}$, but to also be meromorphic at ∞ forces f to be *rational*, i.e. $f(w) = \frac{\text{some polynomial } P(w)}{\text{some polynomial } Q(w)}$, where w is the complex parameter on the Riemann sphere. So our modular function $f(\tau)$ will simply be some rational function P/Q evaluated at the change-of-variables function $w = c(\tau)$ which maps us from our sphere $\Gamma\backslash\overline{\mathcal{H}}$ to the Riemann sphere. There are many different choices for this function $c(\tau)$, but the standard one is $c(\tau) = j(\tau)$, the j-function of (2.1.1)[29]. Thus, any modular function can be written as a rational function $f(\tau) = P(j(\tau))/Q(j(\tau))$ in the j-function. Conversely, any such function will be modular.

This is analogous to saying that any function $g(x)$ periodic under $x \mapsto x+1$ can be thought of as a function on the unit circle $S^1 \subset \mathbb{C}$ evaluated at the change-of-variables function $x \mapsto e^{2\pi \mathrm{i}x}$, and hence has a Fourier expansion $\sum_n g_n \exp[2\pi \mathrm{i}nx]$.

---

[29] Historically, j was the standard choice, but in Moonshine the preferred choice would be the function $J = j - 744$ with zero constant term.

We can generalise this argument. Consider a subgroup G of $\mathrm{SL}_2(\mathbb{R})$ which is both not too big, and not too small. 'Not too big' means it should be *discrete*, i.e. the matrices in G can only get so close to the identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. To make sure G is 'not too small', it is enough to require that G contains some subgroup of the form

$$\Gamma_0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,|\, c \equiv 0 \pmod{N} \} \ , \tag{2.3.2a}$$

i.e. G must contain all matrices in Γ whose bottom-left entry is a multiple of N. So G must contain T, for example. We will also be interested only in those G which obey

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in G \ \Rightarrow \ t \in \mathbb{Z} \tag{2.3.2b}$$

i.e. the only translations in G are by integers. We will call a function $f : \overline{\mathcal{H}} \to \mathbb{C}$ a *modular function for* G if it is meromorphic (including at the cusps $\mathbb{Q} \cup \{\mathrm{i}\infty\}$), and if also f is symmetric with respect to G: $f \circ \alpha = f$ for all $\alpha \in G$. This implies we will be able to expand f as a Laurent series in q. We analyse this as before: look at the orbit space $\Sigma = G \backslash \overline{\mathcal{H}}$; because G is not too big, Σ will be a (Riemann) surface; because G is not too small, Σ will be compact.

The compact Riemann surfaces have been classified (up to homeomorphism — i.e. considering only topology as relevant), and are characterised by a number called the *genus*. Genus 0 is a sphere, genus 1 is a torus, genus 2 is like two tori resting side-by-side, etc. For example, the surface of a wine glass, or a fork, is topologically a sphere, while a coffee cup and a key will (usually) be tori. Eye glasses with the lenses popped out is a 2-torus, while a ladder with n rungs on it has genus $n-1$.

We will call G 'genus g' if its surface Σ has genus g. For example, $G = \Gamma_0(2)$ and $G = \Gamma_0(25)$ are both genus 0, while $\Gamma_0(50)$ is genus 2 and $\Gamma_0(24)$ is genus 3. Once again, we are interested here in the genus 0 case. As before, this means that there is a change-of-variables function we'll denote J_G which has the property that it's a modular function for G, and all other modular functions for G can be written as a rational function in it. Because of (2.3.2), we can choose J_G to look like

$$J_G(\tau) = q^{-1} + a_1(G)\, q + a_2(G)\, q^2 + \cdots$$

So J_G plays exactly the same role for G that $J := j - 744$ plays for Γ. J_G is called the Hauptmodul for G. (Incidentally for genus > 0, two generators, not one, are needed.)

For example, $\Gamma_0(2)$, $\Gamma_0(13)$ and $\Gamma_0(25)$ are all genus 0, with Hauptmoduls

$$J_2(\tau) = q^{-1} + 276\, q - 2048\, q^2 + 11202\, q^3 - 49152\, q^4 + 184024\, q^5 + \cdots \tag{2.3.3}$$

$$J_{13}(\tau) = q^{-1} - q + 2\, q^2 + q^3 + 2\, q^4 - 2\, q^5 - 2\, q^7 - 2\, q^8 + q^9 + \cdots \tag{2.3.4}$$

$$J_{25}(\tau) = q^{-1} - q + q^4 + q^6 - q^{11} - q^{14} + q^{21} + \cdots \tag{2.3.5}$$

The smaller the modular group, the smaller the coefficients of the Hauptmodul. In this sense, the j-function is optimally bad among the Hauptmoduls: e.g. for it $a_{23} \approx 10^{25}$.

An obvious question is, how many genus 0 groups (equivalently, how many Hauptmoduls) are there? It turns out that $\Gamma_0(p)$ is genus 0, for a prime p, iff $p-1$ divides 24. Thompson in 1980 proved that for any g, there are only finitely many genus g groups obeying our two conditions (2.3.2). In particular this means there are only finitely many Hauptmoduls. Over 600 Hauptmoduls with integer coefficients $a_i(G)$ are presently known.

2.4. The Monstrous Moonshine Conjectures

We are now ready to make precise the main conjecture of Conway and Norton [12]. (We should emphasise though that there have been several other conjectures, some of which turned out to be partially wrong.)

They conjectured that for each element g of the Monster $\mathbb{M}$, there is a Hauptmodul

$$J_g(\tau) = q^{-1} + \sum_{n=1}^{\infty} a_n(g)\, q^n \tag{2.4.1}$$

for a genus 0 group G_g such that each coefficient $a_n(g)$ is an integer, and for each n the map $g \mapsto a_n(g)$ is a character of $\mathbb{M}$. They also conjectured that G_g contains $\Gamma_0(N)$ as a normal subgroup, for some N depending on the order of g.

Another way of saying this is that there exists an infinite-dimensional graded representation $V = V_{-1} \oplus \bigoplus_{n=1}^{\infty} V_n$ of $\mathbb{M}$ such that the McKay-Thompson series $T_g(\tau)$ in (2.1.4) is a Hauptmodul.

There are around 8×10^{53} elements to the Monster, so naively we may expect around 8×10^{53} different Hauptmoduls $J_g = T_g$. However the character of a representation evaluated at g and at hgh^{-1} will always be the same, so $J_g = J_{hgh^{-1}}$. Hence the relevant quantity is the number of conjugacy classes, which for $\mathbb{M}$ is only 194. Moreover, a character evaluated at g^{-1} will always be the complex conjugate of its value at g, but here all character values $\chi_{V_n}(g)$ are integers (according to the conjecture). Thus $J_g = J_{g^{-1}}$. The total number of distinct Hauptmoduls J_g arising in Monstrous Moonshine turns out to be only 171.

For example, if we choose g to be the identity, we recover $T_{id.} = J$. It turns out that there are precisely 2 different conjugacy classes of order 2 elements, one of them giving the Hauptmodul J_2 in (2.3.3). Similarly for 13, but J_{25} doesn't correspond to any conjugacy class of $\mathbb{M}$.

Moonshine provides an explanation for a forgotten mystery of classical mathematics: why are the coefficients of the j-function *positive* integers? On the other hand, that they are *integers* has long been important to number theory (complex multiplication, class field theory — see e.g. [16]).

There are lots of other less important conjectures. One which played a role in ultimately proving the main conjecture involves the *replication formulae.* Conway-Norton want to think of the Hauptmoduls J_g as being intimately connected with $\mathbb{M}$; if so, then the group structure of $\mathbb{M}$ should somehow directly relate different J_g. In particular, consider the power map $g \mapsto g^p$. Now, it was well-known that $j(\tau)$ has the property that $j(p\tau) + j(\frac{\tau}{p}) + j(\frac{\tau+1}{p}) + \cdots + j(\frac{\tau+p-1}{p})$ equals a polynomial in j, for any prime p (*sketch* *of proof:* it's a modular function for Γ, and hence equals a rational function of j; since its only poles will be at the cusps, the denominator polynomial must be trivial). Hence the same will hold for J. Explicitly we get

$$J(2\tau) + J(\frac{\tau}{2}) + J(\frac{\tau+1}{2}) = J^2(\tau) - 2a_1 \tag{2.4.2a}$$

$$J(3\tau) + J(\frac{\tau}{3}) + J(\frac{\tau+1}{3}) + J(\frac{\tau+2}{3}) = J^3(\tau) - 3a_1 J(\tau) - 3a_2 \tag{2.4.2b}$$

where $J(\tau) = \sum_k a_k q^k$. Slightly more complicated formulas hold in fact for any composite n. Conway and Norton conjectured that these formulas have an analogue for the Moonshine functions J_g in (2.4.1). In particular, (2.4.2) become for any $g \in \mathbb{M}$

$$J_{g^2}(2\tau) + J_g(\frac{\tau}{2}) + J_g(\frac{\tau+1}{2}) = J_g^2(\tau) - 2a_1(g) \tag{2.4.3a}$$

$$J_{g^3}(3\tau) + J_g(\frac{\tau}{3}) + J_g(\frac{\tau+1}{3}) + J_g(\frac{\tau+2}{3}) = J_g^3(\tau) - 3a_1(g)\, J_g(\tau) - 3a_2(g). \tag{2.4.3b}$$

These are examples of the replication formulae.

'Replication' concerns the power map $g \mapsto g^n$ in $\mathbb{M}$. Can Moonshine see more of the group structure of $\mathbb{M}$? A step in this direction was made by Norton [47], who associated a Hauptmodul to commuting elements g, h in $\mathbb{M}$. Physically [19], this corresponds to orbifold traces, i.e. the $V^\natural$ RCFT with boundary conditions twisted by g and h in the 'time' and 'space' directions. Still, we would like to see more of $\mathbb{M}$ in Moonshine.

An important part of the Monstrous Moonshine conjectures came a few years after [12]. Frenkel-Lepowsky-Meurman [23] constructed a graded infinite-dimensional representation $V^\natural$ of $\mathbb{M}$ and conjectured (correctly) that it is the representation in (2.1.4). $V^\natural$ has a very rich algebraic structure, which will be discussed in §2.6.

A major claim of [23] was that $V^\natural$ is a 'natural' structure (hence their notation). To see what they mean by that, it's best to view another simpler example of a natural construction: that of the *Leech lattice* Λ_{24}. Recall the discussion of (root) lattices in §1.6.

Λ_{24} is one of the most interesting lattices, and is related to Moonshine. It can be defined using 'laminated lattices'. Start with the 0-dimensional lattice $\Lambda_0 = \{0\}$, which consists of just a single point. Use it to construct a 1-dimensional lattice, with minimal (nonzero) norm 4, built out of infinitely many copies of Λ_0 laid side by side. The result of course is simply the even integers $2\mathbb{Z}$, which we will call here Λ_1. Now construct a 2-dimensional lattice, of minimum norm 4, built out of infinitely many copies of Λ_1 laid next to each other. There are lots of ways to do this, but choose the densest lattice possible. The result is unique: it is the hexagonal lattice A_2 scaled by a factor of $\sqrt{2}$: call it Λ_2. Continue in this way: Λ_3, Λ_4, Λ_5, Λ_6, Λ_7, and Λ_8 will be the root lattices A_3, D_4, D_5, E_6, E_7 and E_8, respectively, all scaled by $\sqrt{2}$. See [13] chapter 6 for a more complete treatment of laminated lattices.

The 24th repetition of this construction yields the Leech lattice. It is the unique 24-dimensional self-dual lattice with no norm-2 vectors, and provides among other things the densest known packing of 23-dimensional spheres in $\mathbb{R}^{24}$. Many of its properties are discussed throughout [13]. So lamination provides us with a sort of no-input construction of the Leech lattice, and a good example of the mathematical meaning of 'natural'. After dimension 24, it seems chaos results from the lamination procedure (there are 23 different 25-dimensional lattices that have an equal right to be called Λ_{25}, and over 75 000 are expected for Λ_{26}).

It is natural to ask about Moonshine for other groups. There is a partial Moonshine for the Mathieu groups M_{24} and M_{12} (which have about 2×10^8 and 10^5 elements resp.), the automorphism group .0 of Λ_{24} (which has about 8×10^{18} elements), and a few others — see e.g. [49]. These groups are either simple or almost simple (e.g. .0 is the direct product of $\mathbb{Z}_2$ with the simple group .1). More generally, there will be some sort of Moonshine for any group which is the automorphism group of a vertex operator algebra; the finite simple groups of Lie type should be automorphism groups of VOAs closely related to the affine algebras except defined over fields like $\mathbb{Z}_p$.

There is a geometric side to Moonshine, associated to names like Lian-Yau and Hirzebruch. In particular, Hirzebruch's 'prize question' asks for the construction of a 24-dimensional manifold on which $\mathbb{M}$ acts, whose twisted elliptic genus are the McKay-Thompson series. This is still open.

It should be emphasised that Monstrous Moonshine is a completely unexpected connection between finite groups and modular functions. Although there has been enormous progress in our understanding of this connection (so much so, that Richard Borcherds won the 1998 Fields medal for his work on this), there still is mystery at its heart. In particular, that $\mathbb{M}$ is associated with *modular functions* can be explained mathematically by it being the automorphism group of the Moonshine VOA $V^\natural$, and physically by the associated RCFT, but what is so special about $\mathbb{M}$ that these modular functions should be genus 0? We will come back to this in §2.9.

2.5. FORMAL POWER SERIES

Vertex (operator) algebras (VOAs) are a mathematically precise formulation of the notion of *W-algebra* or *chiral algebra*[30] which is so central to conformal field theory (see §1.1). VOAs were first defined by Borcherds, and their theory has since been developed by a number of people (Frenkel, Lepowsky, Meurman, Zhu, Dong, Li, Mason, Huang, ...). Because our primary motivation here is Moonshine, I will only focus on one aspect of their theory (the connection with Lie algebras). Useful to consult while reading this review are the notes [27] — they take a more analytic approach to many of the things we discuss, and their approach (namely that of CFT) motivates beautifully much of VOA theory.

In quantum field theory the basic object is the quantum field, which roughly speaking is a choice of operator $\hat{A}(x)$ at each space-time point x. 'Operator' means something that 'operates on' functions or vectors. E.g. an indefinite integral is an operator, as is a derivative. The operators in the QFT act on the space spanned by the states $|\star\rangle$, and together form an infinite-dimensional vector space (e.g. a C^* algebra) — this infinite-dimensionality of QFT is a major source of its mathematical difficulties, and QFT still has not been put on completely satisfactory mathematical grounds.

---

[30] An alternate (and much more complicated) mathematical formulation of chiral algebra is due to Beilinson and Drinfeld, and belongs to algebraic geometry. See [28] for a good — but still difficult — review.

But another difficulty is that the quantum field $\hat{A}$ really isn't an operator-valued *function* of space-time. 'Function' is too narrow a concept. For example, one of the most familiar 'functions' in quantum mechanics is the Dirac delta $\delta(x)$. You see it for example in the canonical commutation relations: e.g. for a scalar field $\hat{\varphi}$, we have $[\hat{\varphi}(\vec{x},t), \frac{\partial}{\partial t}\hat{\varphi}(\vec{y},t)] = \mathrm{i}\hbar\delta^3(\vec{x}-\vec{y})$. $\delta(x)$ has the property that for any other smooth function f,

$$\int_{-1}^{1} f(y)\,\delta(y)\,\mathrm{d}y = f(0)\ , \qquad \int_{-1}^{1} f(y)\,\delta'(y)\,\mathrm{d}y = f'(0)\ ,$$

etc. The problem is that $\delta(x)$ isn't a function — no function could possibly have those properties.

One way to make sense of 'functions' like the Dirac delta and its derivatives is distribution theory. Although it was first informally used in physics, it was rigourously developed around 1950 by Laurent Schwartz, and uses the idea of test functions. See e.g. [15].

What I will describe now is an alternate approach, *algebraic* as opposed to *analytic*. These two approaches are not equivalent: you can do some things in one approach which you can't do in the other. But the algebraic approach is considerably simpler technically — no calculus or convergence to worry about — and it is remarkable how much can still be captured. This approach is the starting point for the VOA story described next section, and was first created around 1980 by Garland and Date-Kashiwara-Miwa. Keep in mind that what we are trying to capture is an operator-valued 'function' on space-time. Space-time in CFT is 2-dimensional, and so we can think of it (at least locally) as being on the complex plane $\mathbb{C}$ (more precisely, we will usually associate the space-time point (x,t) with the complex number $z = e^{t+\mathrm{i}x}$). Good introductions to the material in this section are [23,39,31].

Let W be any vector space. We are most interested in it being an infinite-dimensional space of matrices (i.e. operators on an infinite-dimensional space), but forget that for now. Define $W[[z,z^{-1}]]$ to be the set of all formal series $\sum_{n=-\infty}^{\infty} w_n z^n$, where the coefficients w_n lie in our space W. We don't ask here whether a given series converges or diverges — z is merely a formal variable. We will also be interested in $W[z,z^{-1}]$ (Laurent polynomials). We can add these formal series in the usual way, and multiply them by numbers (scalars) in the usual way.

Remember our ultimate aim here: we want to capture quantum fields. So we want our formal series to be operator-valued. The way to accomplish this is to choose W to be a vector space of operators, or matrices if you prefer. A fancy way to say this is '$W = \mathrm{End}(V)$', which means the things in W operate on vectors in V. If we take $V = \mathbb{C}^m$, then we can think of W as being the space of all $m \times m$ complex matrices. We are ultimately interested in the case $m = \infty$, but we won't lose much now by taking $m = 1$, which would mean formal power series with numerical coefficients.

Because our coefficients w_n are operators, we can multiply our formal series. We define multiplication in the usual way. For example, consider $W = V = \mathbb{C}$, and take $c(z) = z^{21} - 5z^{100}$ and $d(z) = \sum_{n=-\infty}^{\infty} z^n$. Then

$$c(z)\,d(z) = \sum_{n=-\infty}^{\infty} z^{n+21} - 5\sum_{n=-\infty}^{\infty} z^{n+100} = \sum_{n=-\infty}^{\infty} z^n - 5\sum_{n=-\infty}^{\infty} z^n = -4d(z)\ .$$

So far so good. Now try to compute the square $d(z)^2$. You get infinity. So the lesson is: you can't always multiply in $W[[z, z^{-1}]]$. We'll come back to this later.

But first, look again at that first product: $c(z)\,d(z) = -4d(z)$. One thing it tells us is that *we can't always divide* (certainly $c(z)$ and -4 are two very different power series!). But there's another lesson here: if you work out a few more multiplications of this kind, what you'll find is that $f(z)\,d(z) = f(1)\,d(z)$ for any f, at least for those f for which $f(1)$ exists (e.g. any $f \in W[z, z^{-1}]$). Thus $d(z)$ is what we would call the Dirac delta $\delta(z-1)$! (You can think of it as the Fourier expansion of the Dirac delta, followed by a change of variables). Unfortunately, the standard notation here is to write it without the '-1':

$$\delta(z) := \sum_{n=-\infty}^{\infty} z^n$$

and that is the notation we will also adopt. Similarly, $\delta(az)$ and $\delta'(z)$ etc (which are the formal series defined in the obvious way) act on $W[z, z^{-1}]$ in the way one would expect: $f(z)\,\delta(az) = f(\frac{1}{a})\,\delta(az)$ and $f(z)\,\delta'(z) = f'(1)\,\delta'(z)$. So of course it makes perfect sense that we couldn't work out $d(z)^2$: we were trying to square the Dirac delta, which we know is impossible!

A similar theory can be developed for several variables z_i, with identities such as $f(z_1, z_2)\,\delta(z_1/z_2) = f(z_2, z_2)\,\delta(z_1/z_2) = f(z_1, z_1)\,\delta(z_1/z_2)$.

But we must not get too overconfident:

Paradox 1. Consider the following product:

$$\delta(z) = \big[(\sum_{n=0}^{\infty} z^n)\,(1-z)\big]\,\delta(z) = (\sum_{n=0}^{\infty} z^n)\,\big[(1-z)\,\delta(z)\big] = (\sum_{n=0}^{\infty} z^n)\,\big[0\,\delta(z)\big] = 0\ .$$

When physicists are confronted with 'paradoxes' such as this, they tend to respond by keeping them in the back of their mind, by treading with care when they are involved in a calculation which reminds them of one of the paradoxes, and otherwise trusting their instincts. Mathematicians typically over-react: they kick themselves for getting overconfident and walking head-first into a 'paradox', and then they devise some rule which will absolutely guarantee that that paradox will always be safely avoided in the future. We will follow the mathematicians' approach, and in the next few paragraphs will describe their rule for avoiding Paradox 1: to forbid certain innocent-looking products.

Remember that we are actually interested in the vector space $W = \mathrm{End}(V)$. Suppose we have infinitely many matrices $w_i \in \mathrm{End}(V)$. We will call them *summable* if for every column vector $v \in V$, only finitely many products $w_i(v) \in V$ are different from 0. In other words, only finitely many of the matrices w_i have a nonzero first column, only finitely many have a nonzero second column, ...

We will certainly have a well-defined sum $\sum_i w_i(z)$ if for each fixed n, the set $\{w_i(n)\}$ (as i varies) of matrices is summable. All other sums are forbidden. We will certainly have a well-defined product[31] $\prod_{i=1}^m w_i(z)$ if for each n, the set $\{w_1(n_1)\,w_2(n_2)\,\cdots\,w_m(n_m)\}_{\sum n_i = n}$

---

[31] m here will be finite: we permit infinite sums but only finite products.

(vary the n_i subject to $\sum_i n_i = n$) is summable. All other products are forbidden. This is reasonable because the sum of those matrix products $w_1(n_1)\cdots w_m(n_m)$ will precisely equal the nth coefficient of the product $\prod_{i=1}^m w_i(z)$.

Note that there are certainly more general ways to have a well-defined product (or sum). For example, according to our rule, we cannot even add $\sum_n 2^{-n}$! This way has the advantage of not touching the more complicated realm of convergence issues. We are doing *algebra* here, not *analysis*. The way out of Paradox 1 is that $(\sum z^n)(1-z)$ doesn't equal 1 — rather, it's a forbidden product.

An interesting consequence of the fact that we are doing algebra instead of analysis is that the product $z^{\frac{1}{2}}\delta(z)$ here does not and cannot equal $1^{\frac{1}{2}}\delta(z) = \delta(z)$ — their formal power series are very different. In hindsight this 'failing' is understandable: algebraically, it seems artificial to prefer the positive root of 1 over the negative root.

Paradox 2. Expand $\frac{1}{1-z}$ in a formal power series in z to get $\sum_{n\geq 0} z^n$. Next, expand $\frac{1}{1-z} = \frac{-z^{-1}}{1-z^{-1}}$ in a formal power series in z^{-1} to get $-\sum_{n<0} z^n$. Subtract these; we presumably should get 0, but we actually get $\delta(z)$!

The analytic explanation is that the first expression converges only for $|z| < 1$, and the second for $|z| > 1$, so it would be naive to expect their difference to be 0. We see from this 'paradox' that *it really matters in which variable we expand rational functions.* For instance, at first glance the identity

$$z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) = z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right)$$

is nonsense; it only holds if you expand the terms in positive powers of z_2, z_1, and z_0 respectively. The procedure of expanding a function in positive and negative powers of a variable and then subtracting the results, yields what are called *expansions of zero*; it is possible to show that expansions of zero will always be linear combinations of Dirac deltas $\delta(az)$ and their various derivatives $\delta^{(k)}(az)$, as we saw in Paradox 2.

2.6. INGREDIENT #3: VERTEX OPERATOR ALGEBRAS

We are now prepared to introduce the important new structure called vertex operator algebras (VOAs). They are essentially the chiral algebras of RCFTs — see [26,27] for excellent motivation of the 7 axioms below. A more detailed treatment of the basic theory of VOAs is provided by e.g. [23,39,31]. Although VOAs are natural from the CFT perspective and appear to be an important and rapidly developing area in math, their definition is not easy: Borcherds is known to have said that you either know what they are, or you don't want to know.

A VOA is a (infinite-dimensional) graded vector space $V = \oplus_{n\in\mathbb{Z}} V_n$ with infinitely many bilinear products $u *_n v$ respecting the grading (in particular $V_k *_n V_\ell \subseteq V_{k+\ell-n-1}$), which obey infinitely many constraints. 'Bilinear' means that for any $a, a', b, b' \in \mathbb{C}$ and $u, u', v, v' \in V$, $(au + a'u') *_n (bv + b'v') = ab\, u *_n v + ab'\, u *_n v' + a'b\, u' *_n v + a'b'\, u' *_n v'$ — i.e. that the products are compatible with the vector space structure of V. The subspaces V_n

must all be finite-dimensional, and they must be trivial (i.e. $V_n = \{0\}$) for all sufficiently small n (i.e. for $n \approx -\infty$). Note that we can collect all these products into one generating function: a linear map $Y : V \to (\mathrm{End}V)[[z, z^{-1}]]$. That is, to each vector $u \in V$ we associate the formal power series (called a *vertex operator*) $Y(u, z) = \sum_{n\in\mathbb{Z}} u_n z^{-n-1}$. For each u, the coefficients u_n will be functions from V to V. The idea is that the product $u *_n v$ will now be written $u_n v := u_n(v)$. The bilinearity of $*_n$ translates into two things in this new language: that $Y(\star, z)$ is linear, and that each function u_n is itself linear (i.e. they are endomorphisms).

The constraints are:

VOA 1. (*regularity*) $u_n v = 0$ for all $n > N(u, v)$;

VOA 2. (*vacuum*) there is a vector $1 \in V$ such that $Y(1, z)$ is the identity (i.e. $1_n v = \delta_{n,-1} v$);

VOA 3. (*state-field correspondence*) $Y(u, 0)1 = u$;

VOA 4. (*conformal*) there is a vector $\omega \in V$, called the *conformal vector*, such that $L_n := \omega_{n+1}$ gives us a representation of the Virasoro algebra $\mathcal{V}$, with central term $C \mapsto cI$ for some $c \in \mathbb{C}$;

VOA 5. (*translation generator*) $Y(L_{-1}u, z) = \frac{d}{dz} Y(u, z)$;

VOA 6. (*conformal weight*) $L_0 u = nu$ whenever $u \in V_n$;

VOA 7. (*locality*) $(z - w)^M [Y(u, z), Y(v, w)] = 0$ for some integer $M = M(u, v)$.

We saw the Virasoro algebra in Part 1 (see (1.2.7)). The number c in **VOA 4** is called the *central charge*, and is an important invariant of V. The peculiar-looking **VOA 7** simply says that the commutator $[Y(u, z), Y(v, w)]$ of two vertex operators will be a finite linear combination of derivatives of various orders of the Dirac delta centred at $z - w$. A recommended exercise for the reader is to show that $M = 4$ works in **VOA 7** for $u = v = \omega$. Note that in a VOA, any $Y(u, z)v$ will be a finite sum — i.e. the series $Y(u, z)$ is summable (defined last section). It is a consequence of the axioms that $1 \in V_0$ and $\omega \in V_2$: for instance, **VOA 7** says all $u_n 1 = 0$ for any $n \geq 0$, so $L_0 1 = \omega_1 1 = 0$ and hence $1 \in V_0$.

In RCFT, V would be the 'Hilbert space of states' (more carefully, V will be a dense subspace of it), and $z = e^{t+ix}$ would be a local complex coordinate on a Riemann surface. L_0 generates time translations, and so its eigenvalues (the *conformal weights*) can be identified with energy. Physically, the requirement that $V_n \to 0$ for $n \to -\infty$ corresponds to the energy of the RCFT being bounded from below. Also, $z = 0$ in **VOA 3** corresponds to the time limit $t \to -\infty$. For each state u, the vertex operator $Y(u, z)$ is a holomorphic (chiral) quantum field. The vector 1 is the vacuum $|0\rangle$, and $Y(\omega, z)$ is the stress-energy tensor T. The most important axiom, **VOA 7**, says that vertex operators commute up to a possible pole at $z = w$, and so are local quantum fields. It is equivalent to the duality axiom of many treatments of CFT. In the physics literature, there is a minor notational difference: for $u \in V_k$, $Y(u, z) = \sum u_n z^{-n-1}$ is written $\sum u_{(n)} z^{-n-k}$. (Physicists prefer this because it cleans up some formulas a little; mathematicians abhor it because it artificially prefers the 'homogeneous' vectors $u \in V_k$.)

In Segal's language (see §1.1), $Y(u, z)$ appears quite naturally. Consider the physical event of two strings combining to form a third. To first order (i.e. the tree-level Feynman diagram), this would correspond in Segal's language to a 'pair-of-pants', or a sphere with three punctures, two of which are positively oriented (corresponding to the incoming strings) and the other being negatively oriented. We can think of the sphere as the Riemann sphere $\mathbb{C}\cup\{\infty\}$; put the punctures at ∞ (outgoing) and z and 0 (the incomings). Segal's functor $\mathcal{T}$ will associate to this a z-dependent homomorphism $\varphi_z : V\times V\to V$. We write $\varphi_z(u,v)\in V$ as $Y(u,z)v$. Incidentally, the symbol 'Y' was chosen because of this 'pair-of-pants' picture (time flows from the top of the 'Y' to the bottom), as was the name 'vertex operator'.

The original axioms by Borcherds were a little more complicated and general: he didn't require $\dim(V_n)<\infty$ nor the $V_n\to 0$ condition, and he only considered L_0 and L_{-1} rather than the full Virasoro algebra. The resulting generalisation is called a *vertex algebra.*

VOA 7 can be rewritten in the form (usually called the *Jacobi identity* for the VOA)

$$z_0^{-1}\delta(\frac{z_1-z_2}{z_0})\,Y(u,z_1)\,Y(v,z_2)-z_0^{-1}\delta(\frac{z_2-z_1}{-z_0})\,Y(v,z_2)\,Y(u,z_1) \tag{2.6.1}$$
$$=z_2^{-1}\delta(\frac{z_1-z_0}{z_2})\,Y(Y(u,z_0)v,z_2)\ ,$$

where the formal series are expanded in the appropriate way. This is the embodiment of commutativity and associativity in the VOA, as we will see. To bring it into a more useful form, hit it with $t\in V$ and expand out into $z_0^\ell z_1^m z_2^n$: we obtain

$$\sum_{i\geq 0}(-1)^i\binom{\ell}{i}\,(u_{\ell+m-i}\circ v_{n+i}-(-1)^\ell v_{\ell+n-i}\circ u_{m+i})=\sum_{i\geq 0}\binom{m}{i}\,(u_{\ell+i}v)_{m+n-i}\ , \tag{2.6.2}$$

where for any $k\in\mathbb{Z}$, $j\in\mathbb{Z}_{\geq}$, $\binom{k}{j}:=\frac{k\,(k-1)\cdots(k-j+1)}{j!}$. For instance, specialising (2.6.2) to $\ell=0$ and $m=0$, resp., gives us

$$[u_m,v_n]=\sum_{i\geq 0}\binom{m}{i}\,(u_iv)_{m+n-i} \tag{2.6.3}$$

$$(u_\ell v)_n=\sum_{i\geq 0}(-1)^i\binom{\ell}{i}\,(u_{\ell-i}\circ v_{n+i}-(-1)^\ell v_{\ell+n-i}\circ u_i)\ . \tag{2.6.4}$$

Why is (2.6.1) called the Jacobi identity? Put $\ell=m=n=0$ in (2.6.2): we get $u_0(v_0t)-v_0(u_0t)=(u_0v)_0t$. If we now formally write $[xy]:=x_0y$, then this becomes $[u[vt]]-[v[ut]]=[[uv]t]$, which is one of the forms of the Lie algebra Jacobi identity (1.2.1b). Even though $[xy]\neq-[yx]$ here, this formal little trick will turn out to be quite important next section.

The simplest examples of VOAs correspond to any even positive-definite lattice Λ; for their construction see e.g. [27,39]. Physically, they correspond to a bosonic string compactified on the torus $\mathbb{R}^n/\Lambda\cong S^1\times\cdots\times S^1$ (where n is the dimension of Λ); the central charge $c=n$. Other important examples, first constructed by Frenkel-Zhu (again see e.g. [27,39]), correspond to affine Kac-Moody algebras $X_\ell^{(1)}$ at level $k\in\mathbb{Z}_>$, and physically to WZW theories on simply-connected compact group manifolds. (We discussed affine algebras in §1.4.) These have central charge $c=\frac{k\dim(X_\ell)}{k+h^\vee}$.

In 1984 Frenkel-Lepowsky-Meurman [23] constructed the *Moonshine module* $V^\natural$. It is a VOA with $c = 24$, with $V^\natural = V_0^\natural \oplus V_1^\natural \oplus V_2^\natural \oplus \cdots$, where $V_0^\natural = \mathbb{C}1$ is 1-dimensional, $V_1^\natural = \{0\}$ is trivial, and $V_2^\natural = (\mathbb{C}\omega) \oplus (\text{Griess algebra})$ is $(1 + 196883)$-dimensional. Its automorphism group (=symmetry group) is precisely the Monster $\mathbb{M}$. Each graded piece $V_n^\natural$ is a finite-dimensional representation of $\mathbb{M}$; Borcherds proved that in fact $V^\natural$ is the McKay-Thompson infinite-dimensional representation of $\mathbb{M}$. It can be regarded as the most natural representation of $\mathbb{M}$ — it is rather surprising that important aspects of a finite group need to be studied via an infinite-dimensional representation.

$V^\natural$ has an elegant physical interpretation. First construct the bosonic string on $\mathbb{R}^{24}/\Lambda_{24}$ (recall that Λ_{24} is the Leech lattice). The resulting $c = 24$ VOA has partition function (=graded dimension) $J(\tau) + 24$, but although its graded pieces (at least for $n > 0$) have the right dimensions, they don't carry a natural representation of $\mathbb{M}$ and so can't qualify for the McKay-Thompson representation. To get $V^\natural$, orbifold this Λ_{24} VOA by the order-2 automorphism of Λ_{24} sending $\vec{x} \mapsto -\vec{x}$. $V^\natural$ thus corresponds to a holomorphic $c = 24$ RCFT, and Moonshine is related to physics. Most of Moonshine can be interpreted physically, except perhaps the genus 0 property of the McKay-Thompson series T_g.

There is a formal parallel between e.g. lattices and VOAs. For example, the Leech lattice Λ_{24} and the Moonshine module $V^\natural$ play analogous roles: Λ_{24} is the unique even lattice which (i) is self-dual, (ii) contains no norm 2 vectors, and (iii) has dimension 24; $V^\natural$ is believed to be the unique VOA which (i) possesses only one irreducible representation (namely itself), (ii) contains no conformal weight 1 elements, and (iii) has central charge $c = 24$. Analogies of these kinds are always useful as they suggest new directions to explore, and the history of math blooms with them. The battlecry 'Why invent what can be profitably copied' is not only heard in Hollywood.

We will end this section on a more speculative note. Witten (1986) said that to understand string theory conceptually, we need a new analogue of Riemannian geometry. Huang (1997) has pushed this thought a little further, saying that there is a more classical 'particle-math' and a more modern 'string-math'. According to Huang we have the real numbers (particle physics) vrs the complex numbers (string theory); Lie algebras vrs VOAs; and the representation theory of Lie algebras vrs RCFT, etc. What are the stringy analogues of calculus, ordinary differential equations, Riemannian manifolds, the Atiyah-Singer Index theorem,...? At present these are all unknown. However, Huang suggests that just as we could imagine Moonshine as a mystery which is explained in some way by RCFT, perhaps the stringy version of calculus would similarly explain the mystery of 2-dimensional gravity, stringy ODEs would explain the mystery of infinite-dimensional integrable systems, stringy Riemannian manifolds would help explain the mystery of mirror symmetry, and the stringy index theorem would help explain the elliptic genus.

2.7. INGREDIENT #4: GENERALISED KAC-MOODY ALGEBRAS

In this section we investigate Lie algebras arising from VOAs. These Lie algebras are an interesting generalisation of Kac-Moody algebras. See e.g. [5,6,38 Chapter 11.13, 29].

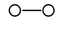
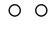
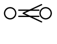
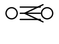
Much of Lie theory (indeed much of algebra) is developed by analogy with simple properties of integers. In §2.2 I invited you to think of a finite group as a massive generalisation of the concept of whole number. Specifically, the number n can be identified with the cyclic group $\mathbb{Z}_n$ with n elements. A divisor d of n generalises to a normal subgroup of a group. A prime number then corresponds to a simple group. Multiplying numbers corresponds to taking the semidirect product of groups (more generally, taking extensions of groups). Then we find that every group has a unique set of simple building blocks (although unlike numbers, different groups can have the same list of building blocks).

For a finite-dimensional Lie algebra, a divisor is called an ideal; a prime is called simple; and multiplying corresponds to semidirect sum. Lie algebras behave simpler than groups but not as simple as numbers, and the analogy sketched above is a reasonably satisfactory one. In particular, simple Lie algebras are important for similar reasons that simple groups are, and as mentioned in §1.2 can also be classified (with *much* less effort). A good treatment of this important classification (over $\mathbb{C}$) is provided by [36]. The proof is now reaching the state of perfection of the formulation of classical mechanics. One unobvious discovery is that the best way to capture the structure of a simple Lie algebra is by an integer matrix, called the *Cartan matrix*, or equivalently but more effectively (since most entries in the Cartan matrix are 0's) by using a graph called the *(Coxeter-)Dynkin diagram.* For instance the Dynkin diagram for A_ℓ consists of ℓ nodes connected sequentially in a line. See Figure 6 in [59].

More precisely, define a *symmetrised Cartan matrix* to be a symmetric real matrix $A = (a_{ij})_{i,j\leq\ell}$ such that $a_{ij} \leq 0$ if $i \neq j$, $a_{ii} > 0$, each $2\frac{a_{ij}}{a_{ii}} \in \mathbb{Z}$, and A is positive-definite. Examples of 2×2 symmetrised Cartan matrices are[32]

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -3 \\ -3 & 6 \end{pmatrix}$$

The Dynkin diagram corresponding to A consists of ℓ nodes; the ith and jth nodes are connected with $4a_{ij}^2/a_{ii}a_{jj}$ lines, and if $a_{ii} \neq a_{jj}$, then we put an arrow over those lines pointing to i if $a_{ii} < a_{jj}$. The Dynkin diagrams corresponding to those four Cartan matrices are respectively

 ,  ,  , 

We may without loss of generality require A to be *indecomposable*, or equivalently that the Dynkin diagram be connected. Of the 4 given above, only the second is decomposable.

To any $\ell \times \ell$ symmetrisable Cartan matrix, we can construct the corresponding Lie algebra $\mathfrak{g}$ in the following way. For each i, create 3 generators e_i, f_i, h_i (so there are a total of 3ℓ generators). The relations these generators obey are given by the following brackets: $[e_i f_j] = \delta_{ij} h_i$, $[h_i e_j] = a_{ij} e_j$, $[h_i f_j] = -a_{ij} f_j$, and for $i \neq j$ $\mathrm{ad}(e_i)^n e_j = \mathrm{ad}(f_i)^n f_j = 0$ where $n = 1 - 2\frac{a_{ij}}{a_{ii}}$. By 'ad$(e)$' here I mean the function $\mathfrak{g} \to \mathfrak{g}$ defined by $\mathrm{ad}(e)f = [ef]$. So $\mathrm{ad}(e)^2 f = [e[ef]]$, $\mathrm{ad}(e)^3 f = [e[e[ef]]]$, etc.

---

[32] Note that our Cartan matrices differ from the usual definition, in which every diagonal entry equals 2.

To get a better feeling for these relations, consider a fixed i. The generators $e = \sqrt{\frac{2}{a_{ii}}}\, e_i, f = \sqrt{\frac{2}{a_{ii}}}\, f_i, h = \frac{2}{a_{ii}}\, h_i$ obey the relations (1.2.2b). In other words, every node in the Dynkin diagram corresponds to a copy of the A_1 Lie algebra. The lines connecting these nodes tells how these ℓ copies of A_1 intertwine.

For instance consider the first Cartan matrix given above. It corresponds to the Lie algebra A_2, or $\mathrm{sl}_3(\mathbb{C})$. The two A_1 subalgebras which generate it (corresponding to the 2 nodes of the Dynkin diagram) can be chosen to be the trace-zero matrices of the form

$$\begin{pmatrix} \star & \star & 0 \\ \star & \star & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & \star & \star \\ 0 & \star & \star \end{pmatrix}.$$

It can be shown that the Lie algebra corresponding to an indecomposable symmetrised Cartan matrix will be finite-dimensional and simple, and conversely that any finite-dimensional simple Lie algebra corresponds to an indecomposable symmetrisable Cartan matrix in this way.

A confusion sometimes arises between the terms 'generators' and 'basis'. Both generators and basis vectors build up the whole algebra; the difference lies in which operations you are permitted to use. For a basis, you are only allowed to use linear combinations (i.e. addition of vectors and multiplication by numbers), while for generators you are also permitted multiplication of vectors (or the bracket, in the Lie case). 'Dimension' refers to basis, while 'rank' usually refers in some way to generators. For instance the (commutative associative) algebra of polynomials in one variable x is infinite-dimensional — any basis needs infinitely many vectors. However, the single polynomial x is enough to generate it (so we could say that its rank is 1). Although those Lie algebras have 3ℓ generators, their dimensions in general will be greater.

From the point of view of generators and relations, the step from 'finite-dimensional simple' to 'symmetrisable Kac-Moody' is rather easy: the only difference is that we drop the 'positive-definite' condition (which was responsible for finite-dimensionality). Kac-Moody (KM) algebras are also generated by (finitely many) A_1 subalgebras, and their theory is quite parallel to that of the simple algebras. Compare Figures 6 and 9 in [59].

Now, it is easy to generalise something; the challenge is to generalise it in a rich and interesting direction. One natural and appealing strategy for generalisation was followed instinctively by a grad student named Robert Moody. Moody's original motivation for developing the theory of Kac-Moody algebras was the Weyl group. If there were Lie algebras for finite Coxeter groups, he asked, why not also for the Euclidean (=affine) ones? For another example of this style of generalisation, consider the question: What is the analogue of calculus (or manifolds) over weird fields — fields (like $\mathbb{Z}_p$) for which the usual limit definitions make no sense? This question leads to the riches of algebraic geometry. Nevertheless this generalisation strategy, even in the hands of a master, will not always be successful. For instance, consider all the trouble the following metaphor has caused: my watch has a maker, so so should the Universe.

Recently Borcherds produced a further generalisation of finite-dimensional simple Lie algebras, which is rather less obvious than that of Kac-Moody algebras. It is easy to associate a Lie algebra to a matrix A, but which class of matrices will yield a deep theory?

Borcherds found such a class by holding in his hand a single algebra (the fake Monster Lie algebra, see §2.9) which acted a lot like a KM algebra, even though it had 'imaginary simple roots'.

By a *generalised symmetrised Cartan matrix* $A = (a_{ij})$ we will mean a symmetric real matrix (possibly infinite), such that $a_{ij} \leq 0$ if $i \neq j$, and if $a_{ii} > 0$ then $2\frac{a_{ij}}{a_{ii}} \in \mathbb{Z}$ for all j. By a *universal generalised Kac-Moody algebra* (universal GKM) or *universal Borcherds-Kac-Moody algebra* $\mathfrak{g}$ we mean the algebra[33] with generators e_i, f_i, h_{ij}, and with relations: $[e_i f_j] = h_{ij}$; $[h_{ij} e_k] = \delta_{ij} a_{ik} e_k$; $[h_{ij} f_k] = -\delta_{ij} a_{ik} f_k$; if $a_{ii} > 0$ and $i \neq j$ then $\mathrm{ad}(e_i)^n e_j = \mathrm{ad}(f_i)^n f_j = 0$ for $n = 1 - 2\frac{a_{ij}}{a_{ii}}$; and if $a_{ij} = 0$ then $[e_i e_j] = [f_i f_j] = 0$.

For example the Heisenberg algebra (1.2.3) corresponds to the choice $A = (0)$, while any other 1×1 $A = (a)$ corresponds to A_1. A universal GKM algebra differs from a KM algebra in that it is built up from Heisenberg algebras as well as A_1, and these subalgebras intertwine in more complicated ways. Nevertheless *much of the theory for finite-dimensional simple Lie algebras continues to find an analogue in this much more general setting* (e.g. root-space decomposition, Weyl group, character formula,...). This unexpected fact is the point of GKM algebras.

To get a feel for these algebras, let us prove a few simple results concerning the h_{ij}. Note first that, using the above relations together with the Jacobi identity, we obtain $[h_{ij} h_{k\ell}] = \delta_{ij}(a_{jk} - a_{j\ell}) h_{k\ell}$. Comparing this with $[h_{k\ell} h_{ij}] = -[h_{ij} h_{k\ell}]$, we see that bracket must always equal 0. Hence all h's pairwise commute, and $h_{ij} = 0$ unless the ith and jth columns of A are identical. An easy exercise now is to show that when $i \neq j$, h_{ij} will lie in the centre of the algebra (i.e. h_{ij} will commute with all other generators).

Although the definition of universal GKM algebra is more natural, it turns out that an equivalent form can be more useful in practice. (It's simpler to describe over $\mathbb{R}$, so in most expositions the reals are used, but alas it's far too late for us to switch loyalties now.) By a *generalised Kac-Moody algebra* (or *Borcherds-Kac-Moody algebra*) $\mathfrak{g}$, we mean a (complex!) Lie algebra which is:

- $\mathbb{Z}$-graded, i.e. $\mathfrak{g} = \oplus_i \mathfrak{g}_i$, where $[\mathfrak{g}_i \mathfrak{g}_j] \subseteq \mathfrak{g}_{i+j}$;
- $\mathfrak{g}_i$ is finite-dimensional for $i \neq 0$;
- $\mathfrak{g}$ has an antilinear involution ω (i.e. $\omega(kx + y) = k^* \omega(x) + \omega(y)$, $[\omega(x)\omega(y)] = \omega([xy])$, and $\omega \circ \omega = id$.) which maps $\mathfrak{g}_i$ to $\mathfrak{g}_{-i}$ and acts as multiplication by -1 on some basis of $\mathfrak{g}_0$;
- $\mathfrak{g}$ has an invariant symmetric bilinear form $(\star, \star)$ (i.e. $([xy], z) = (x, [yz])$ and $(y, z) = (z, y) \in \mathbb{C}$), obeying $(\omega(x), \omega(y)) = (x, y)^*$, such that $(\mathfrak{g}_i, \mathfrak{g}_j) = 0$ if $i \neq -j$;
- the Hermitian form defined by $(x|y) := -(\omega(x), y)$ is positive-definite on $\mathfrak{g}_{i \neq 0}$.

Note that for some basis x_i of $\mathfrak{g}_0$, the third condition tells us $-[x_i x_j] = [(-x_i)(-x_j)] = [x_i x_j]$, i.e. $\mathfrak{g}_0$ has a trivial bracket. It plays the role of the Cartan subalgebra $\mathfrak{h}$ in the theory.

For example, let $\mathfrak{g} = \mathrm{sl}_2(\mathbb{C})$ and recall (1.2.5). Then $\mathfrak{g}_1 = \mathbb{C}e$, $\mathfrak{g}_0 = \mathbb{C}h$, $\mathfrak{g}_{-1} = \mathbb{C}f$ is the root-space decomposition. $\omega(x) = -x^\dagger$ is the Cartan involution. $(x, y) = \mathrm{tr}(xy)$ is the Killing form.

---

[33] As with KM algebras, usually we want to extend it by some derivations; enough derivations are added so that the simple roots are linearly independent.

It turns out [5] that any universal GKM algebra is a GKM algebra, and any GKM algebra can be constructed from a unique universal GKM algebra (by quotienting out part of the centre and adding derivations), so in that sense the two structures are equivalent. This theorem is important, because it tells us that *GKM algebras are the ultimate generalisation of simple Lie algebras,* in the sense that any further generalisation will lose some basic structural ingredient.

We know simple Lie algebras (and groups) arise in both classical and quantum physics, and the affine KM algebras are important in CFT, as we saw in Part 1. GKM algebras have recently appeared in the physics literature (see Harvey-Moore) in the context of BPS states in string theory.

How do GKM algebras arise in VOAs? If we define $[xy] := x_0 y$, then as mentioned in §2.6 we get from the VOA Jacobi identity the equation $[x[yz]] - [y[xz]] = [[xy]z]$. Thus our bracket will be anti-associative if it is anti-commutative. But is it anti-commutative? It can be shown

$$u_n v = \sum_{i=0}^{\infty} \frac{1}{i!}(-1)^{i+n+1}(L_{-1})^i (v_{n+i} u) \tag{2.7.1}$$

so $u_0 v \equiv -v_0 u$ if we look at things mod $L_{-1}V$.

Since our bracket is clearly bilinear, we thus get a Lie algebra structure on $V/L_{-1}V$. Similarly, we get a symmetric bilinear product on $V/L_{-1}V$, given by $\langle u, v\rangle := u_1 v$.

We would like $\langle \star, \star\rangle$ to respect the Lie algebra structure, i.e. be $[\star\star]$-invariant. We compute from (2.6.4)

$$\langle [uv], t\rangle = -[v\langle u, t\rangle] + \langle u, [vt]\rangle \ . \tag{2.7.2}$$

Since we would like to identify $\langle \star, \star\rangle$ with the bilinear form in the GKM algebra definition, we also would like it to be number-valued (i.e. have 1-dimensional range).

There is a simple way to satisfy both of these. First, restrict attention to V_1, i.e. the conformal weight 1 vectors: $V_1 \cap (V/L_{-1}V) = V_1/L_{-1}V_0$. Then $\langle u, v\rangle \in V_0$. Assume V_0 is 1-dimensional: i.e. $V_0 = \mathbb{C}1$. Then $\langle u, v\rangle$ will equal a number times 1, so call (u, v) that number. Also, $[\langle u, t\rangle v] = (u, t)1_0 v = 0$, so $\langle \star, \star\rangle$ and hence $(\star, \star)$, will be invariant. Of course, when $V_0 = \mathbb{C}1$, $L_{-1}V_0 = \{0\}$.

$V_1/L_{-1}V_0$ is generally too large in practice to be useful; a subalgebra can be defined as follows. Let P_n be the 'primary states with conformal weight n', i.e. the $u \in V_n$ killed by L_m for all $m > 0$. Then $\mathfrak{g}(V) := P_1/L_{-1}P_0$ will be a subalgebra of $V_1/L_{-1}V_0$. Through the assignment $u \mapsto u_0$, $\mathfrak{g}(V)$ acts on V and this action commutes with that of L_i. This association of a Lie algebra to a VOA is due to Borcherds (1986).

Similar arguments show that when V_0 is 1-dimensional and V_1 is 0-dimensional, then V_2 will necessarily be a commutative nonassociative algebra with product $u \times v := u_1 v \in V_2$ and identity element $\frac{1}{2}\omega$ (*proof*: $\omega \times u = L_0 u = 2u$). Now, those conditions on V_0, V_1 are satisfied by the Moonshine module $V^\natural$. We find that $V_2^\natural$ is none other than the Griess algebra extended by an identity element.

2.8. INGREDIENT #5: DENOMINATOR IDENTITIES

In §1.3, we discussed the representation theory of Lie algebras. An important invariant of a representation is its *character*. Simple Lie algebras possess a very useful formula for their characters, due to Weyl:

$$\mathrm{ch}_\lambda(z) := \sum_\mu \dim(L_\lambda(\mu))\, e^{\mu\cdot z} = e^{-\rho\cdot z}\, \frac{\sum_{w\in W} \pm e^{w(\lambda+\rho)\cdot z}}{\prod_{\alpha\in\Delta_+}(1-e^{-\alpha\cdot z})}\ , \tag{2.8.1}$$

where W is the Weyl group, Δ_+ the positive roots, and where $\oplus_\mu L_\lambda(\mu)$ is the weight-space decomposition of L_λ — i.e. the simultaneous eigenspaces of the h_i. Here z belongs to the Cartan subalgebra $\mathfrak{h}$; the character is complex-valued. Analogous statements hold for all GKM algebras.

It is rare indeed when a trivial special case of a theorem or formula produces something interesting. But that is what happens here. Consider the trivial representation: i.e. $x \mapsto 0$ for all $x \in X_\ell$. Then the character is identically 1, by definition: $ch_0 \equiv 1$. Thus the character formula tells us that a certain alternating sum over a Weyl group, equals a certain product over positive roots. These formulas, called *denominator formulas*, are nontrivial even in the finite-dimensional cases.

Consider for instance the smallest simple algebra, A_1. Here the identity indeed is too trivial: it reads $e^{z/2} - e^{-z/2} = e^{z/2}(1-e^{-z})$. For A_2 we get a sum of 6 terms equalling a product of 3 terms, and the complexity continues to rise from there.

Around 1970 Macdonald tried to generalise these finite denominator identities to infinite identities, corresponding to the extended Dynkin diagrams. These were later reinterpreted by Kac, Moody and others as denominator identities for affine nontwisted KM algebras. The simplest one was known classically as the Jacobi triple product identity:

$$\sum_{n=-\infty}^{\infty}(-1)^n x^{n^2} y^n = \prod_{m=1}^{\infty}(1-x^{2m})(1-x^{2m-1}y)(1-x^{2m-1}y^{-1})$$

We now know it to be the denominator identity for the simplest infinite-dimensional KM algebra, $A_1^{(1)}$.

Freeman Dyson is famous for his work in quantum field theory, but he started as an undergraduate in number theory and still enjoys it as a hobby. Dyson [20] found a curious formula for the Ramanujan τ-function, which can be defined by the generating function $\sum_{n=1}^\infty \tau(n)x^n = \eta^{24}(x) := x\prod_{m=1}^\infty (1-x^m)^{24}$. Dyson found the remarkable formula

$$\tau(n) = \sum \frac{(a-b)(a-c)(a-d)(a-e)(b-c)(b-d)(b-e)(c-d)(c-e)(d-e)}{1!\ 2!\ 3!\ 4!}$$

where the sum is over all 5-tuples $(a,b,c,d,e) \equiv (1,2,3,4,5) \pmod 5$ obeying $a+b+c+d+e=0$ and $a^2+b^2+c^2+d^2+e^2 = 10n$. Using this, an analogous formula can be found for η^{24}. Dyson knew that similar-looking formulas were also known for η^d for the values $d = 3, 8, 10, 14, 15, 21, 24, 26, 28, 35, 36, \ldots$.

What was ironic was that Dyson found that formula at the same time that Macdonald was finding the Macdonald identities. Both were at Princeton then, and would often chat a little when they bumped into each other after dropping off their daughters at school. But they never discussed work. Dyson didn't realise that his strange list of numbers has a simple interpretation: they are precisely the dimensions of the simple Lie algebras! $3 = \dim(\mathrm{A}_1)$, $8 = \dim(A_2)$, $10 = \dim(C_2)$, $14 = \dim(G_2)$, etc. In fact these formulas for η^d are none other than (specialisations of) the Macdonald identities. For example, Dyson's formula is the denominator formula for $A_4^{(1)}$ ($24 = \dim(A_4)$). If they had spoken, they would probably have anticipated the affine denominator identity interpretation.

One curiousity apparently still has no algebraic interpretation: No simple Lie algebra has dimension 26, so the formula for η^{26} can't correspond to any Macdonald identity.

Macdonald didn't close the book on denominator identities. More recently Kac and Wakimoto [40] have used denominator identities for Lie superalgebras to obtain nice formulas for various generator functions involving sums of squares, sums of triangular numbers (triangular numbers are numbers of the form $\frac{1}{2}k(k+1)$), etc. For instance, the number of ways n can be written as a sum of 16 triangular numbers is

$$\frac{1}{3\cdot 4^3}\sum ab\,(a^2-b^2)^2$$

where the sum is over all odd positive integers a, b, r, s obeying $ar + bs = 2n + 4$ and $a > b$.

Another example of denominator identities is Borcherds' use of them in proving the Moonshine conjectures. In particular this motivated his introduction of the GKM algebras. The denominator identities for other GKM algebras were used by Borcherds to obtain results on the moduli spaces of e.g. families of K3 surfaces. They are also often turned-around now and used for learning about the positive roots in a given GKM.

2.9. PROOF OF THE MOONSHINE CONJECTURES

The main Conway-Norton conjecture was proved almost immediately. Thompson showed that if $g \mapsto a_n(g)$ is a character for all $n \leq 1200$, then it will be for all n. He also showed that if certain congruence conditions hold for a certain number of $a_n(g)$ (all with $n \leq 100$), then all $g \mapsto a_n(g)$ will be *virtual* characters (i.e. a linear combination over $\mathbb{Z}$ of irreducible characters of $\mathbb{M}$; only if all coefficients are nonnegative will it be a true character). Atkin-Fong-Smith [54] used that to prove on a computer that indeed all were virtual characters. But their work didn't say anything about the underlying (possibly virtual) representation V. The real challenge was to construct (preferably in a natural manner) the representation which works. Frenkel-Lepowsky-Meurman [23] constructed a candidate for it (the Moonshine module $V^\natural$); it was Borcherds who finally proved $V^\natural$ obeyed the Conway-Norton conjecture. A good overview of Borcherds' work on Moonshine is provided in [32].

We want to show that the McKay-Thompson series $T_g(\tau) := q^{-1}\mathrm{tr}_{V^\natural}(gq^{L_0})$ of (2.1.4) equals the Hauptmodul $J_g(\tau)$ in (2.4.1) (the 'fudge factor' $q^{-1} = q^{-c/24}$ is familiar to e.g. KM algebras and CFT and was discussed at the end of §1.2). Borcherds' strategy was to bring in Lie theory and to use the corresponding denominator identity to provide useful combinatorial data. The first guess for this 'Monster Lie algebra' was the Kac-Moody algebra whose Dynkin diagram is essentially the Leech lattice (i.e. a node for each vector in Λ_{24}, and 2 nodes are connected by a number of edges depending on the value of a certain dot product). It was eventually discarded because some of the critical data (namely positive root multiplicities) needed in order to write down its denominator identity were too complicated. But looking at that failed attempt led Borcherds to a second candidate, now called the *fake Monster Lie algebra* $\mathfrak{g}'_{\mathbb{M}}$. In order to *construct* it, he developed the theory of VOAs; and in order to *understand* it, he developed the theory of GKM algebras. We will define it shortly. $\mathfrak{g}'_{\mathbb{M}}$ also turned out to be inadequate for proving the Moonshine conjectures; however it directly led him to the GKM algebra now called the true *Monster Lie algebra* $\mathfrak{g}_{\mathbb{M}}$. And that directly led to the proof of Moonshine.

Step 1: Construct $\mathfrak{g}_{\mathbb{M}}$ from $V^\natural = V_0^\natural \oplus V_1^\natural \oplus \cdots$. For later convenience, reparametrise these subspaces $V^i := V_{i+1}^\natural$. Recall the even indefinite lattice $II_{1,1}$ defined in (1.6.1). Of course the direct choice $\mathfrak{g}(V^\natural)$ is 0-dimensional because $V_1^\natural$ is trivial, so we must modify $V^\natural$ first.

The Monster Lie algebra is (essentially) defined to be the Lie algebra $\mathfrak{g}(V^\natural \otimes V_{II_{1,1}})$ associated to the vertex algebra $V^\natural \otimes V_{II_{1,1}}$ (strictly speaking, more of $\mathfrak{g}$ is quotiented away). By contrast, the fake Monster is the Lie algebra associated to the vertex algebra $V_{\Lambda_{24}} \otimes V_{II_{1,1}} \cong V_{II_{25,1}}$. Both of these are vertex algebras as opposed to VOAs, because of the presence of the indefinite lattices, but this isn't important here. $\mathfrak{g}_{\mathbb{M}}$ inherits a $II_{1,1}$-grading from $V_{II_{1,1}}$: the piece of grading (m,n) is isomorphic (as a vector space) to V^{mn}, if $(m,n) \neq (0,0)$; the (0,0) piece is isomorphic to $\mathbb{R}^2$. Borcherds uses the No-Ghost Theorem of string theory to show that the homogeneous pieces of $\mathfrak{g}_{\mathbb{M}}$ are those of $V^\natural$.

Both $\mathfrak{g}_{\mathbb{M}}$ and $\mathfrak{g}'_{\mathbb{M}}$ are GKM algebras; for instance the $\mathbb{Z}$-grading of $\mathfrak{g}_{\mathbb{M}}$ is given by $(\mathfrak{g}_{\mathbb{M}})_k = \oplus_{m+n=k} V^{mn}$ for $k \neq 0$, while the 0-part is $V^{-1} \oplus V^{-1}$. Although $\mathfrak{g}'_{\mathbb{M}}$ is not used in the proof of the Monstrous Moonshine conjectures, it is related to some kind of Moonshine for the finite simple group .1, which is 'half' of the automorphism group .0 of the Leech lattice Λ_{24}. $\mathfrak{g}_{\mathbb{M}}$ corresponds to the Cartan matrix

$$\begin{pmatrix} B_{-1,-1} & B_{-1,1} & B_{-1,2} & \cdots \\ B_{1,-1} & B_{1,1} & B_{1,2} & \cdots \\ B_{2,-1} & B_{2,1} & B_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $B_{i,j}$ for $i,j \in \{-1,1,2,3,\ldots\}$ is the $a_i \times a_j$ block with fixed entry $-i-j$ (the a_i as usual are the coefficients $\sum_n a_n q^n$ of $j - 744$).

2nd step: Compute the denominator identity of $\mathfrak{g}_{\mathbb{M}}$: we get

$$p^{-1} \prod_{\substack{m>0 \\ n\in\mathbb{Z}}} (1 - p^m q^n)^{a_{mn}} = j(z) - j(\tau) \tag{2.9.1}$$

where $p = e^{2\pi \mathrm{i} z}$. The result are various formulas involving the coefficients a_i, for instance $a_4 = a_3 + (a_1^2 - a_1)/2$. It turns out to be possible to 'twist' (2.9.1) by each $g \in \mathbb{M}$, obtaining

$$p^{-1} \exp[-\sum_{k>0} \sum_{\substack{m>0 \\ n\in\mathbb{Z}}} a_{mn}(g^k) \frac{p^{mk} q^{nk}}{k}] = T_g(z) - T_g(\tau)\ . \tag{2.9.2}$$

This looks a lot more complicated, but you can glimpse the Taylor expansion of $\ln(1-p^m q^n)$ there and in fact for $g = id$ (2.9.2) reduces to (2.9.1). This formula gives more generally identities like $a_4(g) = a_2(g) + (a_1(g)^2 - a_1(g^2))/2$, where $T_g(\tau) = \sum_i a_i(g) q^i$. These formulas involving the McKay-Thompson coefficients are equivalent to the replication formulae conjectured in §2.4.

3rd step: It was known earlier that all of the Hauptmoduls also obey those replication formula, and that anything obeying them will lie in a finite-dimensional manifold which we'll call R. In particular, if $B(q) = q^{-1} + \sum_{n>0} b_n q^n$ and $C(q) = q^{-1} + \sum_{n>0} c_n q^n$ both lie in R, and $b_n = c_n$ for $n \leq 23$, then $B(q) = C(q)$. In fact, it turns out that if we verify for each conjugacy class $[g]$ of $\mathbb{M}$ that the first, second, third, fourth and sixth coefficients of the McKay-Thompson series T_g and the corresponding Hauptmodul J_g agree, then $T_g = J_g$, and we are done.

That is precisely what Borcherds then did: he compared finitely many coefficients, and as they all equalled what they should, this concluded the proof of Monstrous Moonshine!

However there was a disappointing side to his proof. While no one disputed its *logical* validity, it did seem to possess a disappointing *conceptual gap*. In particular, the Moonshine conjectures were made in the hope that proving them would help explain what the Monster had to do with the j-function and the other Hauptmoduls. A good proof says much more than 'True' or 'False'. The case-by-case verification occurred at the critical point where the McKay-Thompson series were being compared directly to the Hauptmoduls. The proof showed that indeed the Moonshine module establishes some sort of relation between T_g and J_g (namely, they must lie in the same finite-dimensional space), but why couldn't it be just a happy meaningless accident that they be equal? Of course we believe it's more than merely an accident, so our proof should reflect this: we want a more conceptual explanation.

This conceptual gap has since been filled [17] — i.e. the case-by-case verification has been replaced with a general theorem. It turns out that something obeying the replicable formulas will also obey something called *modular equations*. A modular equation for a function f is a polynomial identity obeyed by $f(x)$ and $f(nx)$. The simplest examples come from the exponential and cosine functions: note that for any $n > 0$, $\exp(nx) = (\exp(x))^n$ and $\cos(nx) = T_n(\cos(x))$ where T_n is a Tchebychev polynomial. A more interesting example of a modular equation is obeyed by $J(\tau) = j(\tau) - 744$: put $X = J(\tau)$ and $Y = J(2\tau)$, then

$$(X^2 - Y)(Y^2 - X) = 393768\,(X^2 + Y^2) + 42987520\,XY + 40491318744\,(X + Y) \\ - 120981708338256\ .$$

Finding modular equations (for various elliptic functions) was a passion of the great mathematician Ramanujan. His notebooks are filled with them. See e.g. [7] for an application of Ramanujan's modular equations to computing the first billion or so digits of π. Many modular equations are also studied in [10]. For more of their applications, see e.g. [16]. It can be shown that the only functions $f(\tau) = q^{-1} + a_1 q + \cdots$ which obey modular equations for all n, are $J(\tau)$ and the 'modular fictions' q^{-1} and $q^{-1} \pm q$ (which are essentially exp, cos, and sin).

It was proved in [17] that, roughly speaking, a function $B(\tau) = q^{-1} + \sum_{n>0} b_n q^n$ which obeys enough modular equations, will either be of the form $B(\tau) = q^{-1} + b_1 q$, or will necessarily be a Hauptmodul for a modular group containing some $\Gamma_0(N)$. The converse is also true: for instance, a modular equation for the Hauptmodul J_{25} of $\Gamma_0(25)$ given in (2.3.5) is

$$(X^2 - Y)(Y^2 - X) = -2\,(X^2 + Y^2) + 4\,(X + Y) - 4\ ,$$

where $X = J_{25}(\tau)$ and $Y = J_{25}(2\tau)$. To eliminate the conceptual gap, this result should then replace step 3. Steps 1 and 2 are still required, however.

This conceptual gap should not take away from what was a remarkable accomplishment by Borcherds: not only the proof of the Monstrous Moonshine conjectures, but also the definition of two new and important algebraic structures. I hope the preceding sections give the reader some indication of why Borcherds was awarded one of the 1998 Fields Medals.

Another approach to the Hauptmodul property is by Tuite [57], who related it to the (conjectured) uniqueness of $V^\natural$. Norton has suggested that the reason $\mathbb{M}$ is associated to *genus-zero* modular functions could be what he calls its '6-transposition' property [47].

So has Moonshine been explained? According to Conway, McKay, and many others, it hasn't. They consider VOAs in general, and $V^\natural$ in particular, to be too complicated to be God-given. The progress, though impressive, has broadened not lessened the fundamental mystery, they would argue.

For what it's worth, I don't completely agree. Explaining away a mystery is a little like grasping a bar of soap in a bathtub, or quenching a child's curiousity. Only extreme measures like pulling the plug, or enrollment in school, ever really work. True progress means displacing the mystery, usually from the particular to the general. Why is the sky blue? Because of how light scatters in gases. Why are Hauptmoduls attached to each $g \in \mathbb{M}$? Because of $V^\natural$. Mystery exists wherever we can ask 'why' — like beauty it's in the beholder's eye.

Moonshine is now 'leaving the nest'. We are entering a consolation phase, tidying up, generalising, simplifying, clarifying, working out more examples. Important and interesting discoveries will be made in the next few years, and yes there still is mystery, but no longer does a Moonshiner feel like an illicit distiller: Moonshine is now a day-job!

Acknowledgments. I warmly thank the Feza Gursey Institute in Istanbul, and in particular Teoman Turgut, for their invitation to the workshop and hospitality during my month-long stay. These notes are based on 16 lectures I gave there in Summer 1998. I've also benefitted from numerous conversations with Y. Billig, A. Coste, C. Cummins, M. Gaberdiel, J. McKay, M. Tuite, and M. Walton — Mark Walton in particular made a very careful reading of the manuscript (and hence must share partial blame for any errors still remaining). My appreciation as well goes to J.-B. Zuber and P. Ruelle for sharing with me their personal stories behind the discoveries of, respectively, A-D-E in $A_1^{(1)}$ and Fermat in $A_2^{(1)}$. The research was supported in part by NSERC.

REFERENCES

1. V. I. Arnold, *Catastrophe Theory*, 2nd edn. (Springer, Berlin, 1997);
   M. Hazewinkel, W. Hesselink, D. Siersma, and F.D. Veldkamp, *Nieuw Arch. Wisk.* **25** (1977) 257;
   P. Slodowy, in: Lecture Notes in Math 1008, J. Dolgachev (ed.) (Springer, Berlin, 1983).
2. M. Bauer, A. Coste, C. Itzykson, and P. Ruelle, *J. Geom. Phys.* **22** (1997) 134.
3. D. Bernard, *Nucl. Phys.* **B288** (1987) 628.
4. J. Böckenhauer and D. E. Evans, *Commun. Math. Phys.* **200** (1999) 57; "Modular invariants, graphs and α-induction for nets of subfactors III", hep-th/9812110.
5. R. E. Borcherds, *Invent. math.* **109** (1992) 405.
6. R. E. Borcherds, "What is moonshine?", math.QA/9809110.
7. J. M. Borwein, P. B. Borwein and D. H. Bailey, *Amer. Math Monthly* **96** (1989) 201.
8. A. Cappelli, C. Itzykson, and J.-B. Zuber, *Commun. Math. Phys.* **113** (1987) 1.
9. R. Carter, G. Segal and I. M. Macdonald, *Lectures on Lie Groups and Lie algebras* (Cambridge University Press, Cambridge, 1995).
10. A. Cayley, *An Elementary Treatise on Elliptic Functions*, 2nd edn (Dover, New York, 1961).
11. J. H. Conway, *Math. Intelligencer* **2** (1980) 165.
12. J. H. Conway and S. P. Norton, *Bull. London Math. Soc.* **11** (1979) 308.
13. J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, 3rd edn (Springer, Berlin, 1999).
14. A. Coste and T. Gannon, *Phys. Lett.* **B323** (1994) 316.
15. R. Courant and D. Hilbert, *Methods of Mathematical Physics II* (Wiley, New York, 1989).
16. D. Cox, *Primes of the form $x^2 + ny^2$*, (Wiley, New York, 1989).
17. C. J. Cummins and T. Gannon, *Invent. math.* **129** (1997) 413.
18. P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory* (Springer, New York, 1996).
19. L. Dixon, P. Ginsparg, and J. Harvey, *Commun. Math. Phys.* **119** (1988) 221.
20. F. Dyson, *Bull. Amer. Math. Soc.* **78** (1972) 635.
21. F. J. Dyson, *Math. Intelligencer* **5** (1983) 47.
22. D.E. Evans and Y. Kawahigashi, *Quantum symmetries on operator algebras* (Oxford University Press, Oxford, 1998).
23. I. Frenkel, J. Lepowsky, and A. Meurman, *Vertex operator algebras and the Monster* (Academic Press, San Diego, 1988).
24. J. Fuchs and C. Schweigert, *Symmetries, Lie algebras, and representations* (Cambridge University Press, Cambridge, 1997).
25. W. Fulton and J. Harris, *Representation Theory: A first course* (Springer, New York, 1996).
26. M. R. Gaberdiel and P. Goddard, "Axiomatic conformal field theory", hep-th/9810019.
27. M. R. Gaberdiel and P. Goddard, "An introduction to meromorphic conformal field theory and its representations", lecture notes in this volume.

28. D. Gaitsgory, "Notes on two dimensional conformal field theory and string theory", math.AG/9811061.
29. T. Gannon, "The Cappelli-Itzykson-Zuber A-D-E classification", math.QA/9902064.
30. K. Gawedzki, "Conformal field theory: a case study", lecture notes in this volume.
31. R. W. Gebert, *Internat. J. Mod. Phys.* **A8** (1993) 5441.
32. P. Goddard, "The work of R.E. Borcherds", math.QA/9808136.
33. D. Gorenstein, *Finite Simple Groups* (Plenum, New York, 1982).
34. D. Gorenstein, R. Lyons, and R. Solomon, *The Classification of the Finite Simple Groups* (AMS, Providence, 1994).
35. A. Hanany and Y.-H. He, "Non-abelian finite gauge theories", hep-th/9811183.
36. J. E. Humphreys, *Introduction to Lie algebras and representation theory* (Springer, New York, 1994).
37. V. G. Kac, In: Lecture Notes in Math 848 (Springer, New York, 1981).
38. V. G. Kac, *Infinite Dimensional Lie algebras*, 3rd edn (Cambridge University Press, Cambridge, 1990).
39. V. G. Kac, *Vertex Algebras for Beginners*, 2nd edn (AMS, Providence, 1998).
40. V. G. Kac and M. Wakimoto, In: *Lie Theory and Geometry in Honor of Bertram Kostant*, Progress in Math. **123** (Birkhäuser, Boston, 1994).
41. S. Kass, R. V. Moody, J. Patera and R. Slansky, *Affine Lie algebras, weight multiplicities, and branching rules*, Vol. 1 (University of California Press, Berkeley, 1990).
42. S. Lang, *Elliptic Functions*, 2nd edn (New York, Springer, 1997).
43. F. W. Lawvere and S. H. Schanuel, *Conceptual Mathematics: A first introduction to Categories* (Cambridge University Press, Cambridge, 1997).
44. H. Minc, *Nonnegative matrices* (Wiley, New York, 1988).
45. E.J. Mlawer, S.G. Naculich, H.A. Riggs, and H. J. Schnitzer, *Nucl. Phys.* **B352** (1991) 863.
46. W. Nahm, *Commun. Math. Phys.* **118** (1988) 171.
47. S. P. Norton, In: *Proc. Symp. Pure Math.* **47** (1987) 208.
48. A. Ocneanu, "Paths on Coxeter diagrams: From Platonic solids and singularities to minimal models and subfactors" (Lectures given at Fields Institute (1995), notes recorded by S. Goto).
49. L. Queen, *Math. of Comput.* **37** (1981) 547.
50. A.N. Schellekens and S. Yankielowicz, *Nucl. Phys.* **B327** (1989) 673.
51. M. Schottenloher, *A Mathematical Introduction to Conformal Field Theory* (Springer, Berlin, 1997).
52. G. Segal, In: *IXth Proc. Int. Congress Math. Phys.* (Hilger, 1989).
53. S. Singh, *Fermat's Enigma* (Penguin Books, London, 1997).
54. S. D. Smith, In: *Finite Groups – Coming of Age*, Contemp. Math. **193** (AMS, Providence, 1996).
55. W. Thurston, *Three-dimensional geometry and topology*, vol. 1 (Princeton, 1997).
56. L. Toti Rigatelli, *Evariste Galois* (Birkhäuser, Basel, 1996).
57. M. Tuite, *Commun. Math. Phys.* **166** (1995) 495.
58. V. G. Turaev, *Quantum invariants of knots and 3-manifolds* (de Gruyter, Berlin, 1994).

59. M. A. Walton, “Affine Kac-Moody algebras and the Wess-Zumino-Witten model”, lecture notes in this volume.
60. M. Waldschmidt, P. Moussa, J.-M. Luck, and C. Itzykson (ed.), *From Number Theory to Physics* (Berlin, Springer, 1992).
61. D. Zagier, *Amer. Math. Monthly* **97** (1990) 144.
62. J.-B. Zuber, *Commun. Math. Phys.* **179** (1996) 265.
63. J.-B. Zuber, private communication, February 1999.

arXiv:math/0402345v2 [math.QA] 14 Apr 2004

Monstrous Moonshine: The first twenty-five years

Terry Gannon

Abstract

Twenty-five years ago, Conway and Norton published in this journal[†] their remarkable paper 'Monstrous Moonshine', proposing a completely unexpected relationship between finite simple groups and modular functions. This paper reviews the progress made in broadening and understanding that relationship.

1. Introduction

It has been approximately twenty-five years since John McKay remarked that

$$196\,884 = 196\,883 + 1\ . \tag{1.1}$$

That time has seen the discovery of important structures, the establishment of another deep connection between number theory and algebra, and a reinforcement of a new era of cooperation between pure mathematics and mathematical physics. It is a beautiful and accessible example of how mathematics can be driven by strictly conceptual concerns, and of how the particular and the general can feed off each other. Now, six years after Borcherds' Fields Medal, the original flurry of activity is over; the new period should be one of consolidation and generalisation and should witness the gradual movement of this still rather esoteric corner of mathematics toward the mainstream.

The central question McKay's equation (1.1) raises, is: What does the j-function (the left side) have to do with the Monster finite group (the right side)? Many would argue that we still don't have our finger on the essence of the matter. But what is clear is that we understand far more about this central question today than we did in 1978. Today we say that there is a vertex operator algebra, called the *Moonshine module* $V^\natural$, which interpolates between the left and right sides of (1.1): its automorphism group is the Monster and its graded dimension is the j-function (-744).

This paper tries to summarise this work of the past twenty-five years in about as many pages. The original article [**24**] is still very readable and contains a wealth of information not found in other sources. Other reviews are [**21**], [**87**], [**12**], [**39**], [**90**], [**48**], [**15**], [**78**], [**102**], [**16**], [**46**] and the introductory chapter in [**44**], and each has its own emphasis. Our own bias here has been to breadth at the expense of depth, which probably limits this review to be a mere annotated sampling of representative literature.

2. Background

In §2.1 we describe the finite simple groups and in particular the Monster. In §2.2 we focus on the modular groups and functions which arise in Monstrous Moonshine.

2.1. The Monster. By definition, a simple group is one whose only normal subgroups are the trivial ones: {1} and the group itself. The importance of the finite simple groups lies in their role as building blocks, in the sense that any finite group G can be constructed from $\{1\}$ by extending successively by a (unique up to order) sequence of finite simple groups. For example the symmetric group S_4 arises in this way from the cyclic groups C_2, C_2, C_3, C_2.

[†] This is an invited paper for the *Bull. London Math. Soc.* Comments are welcome.

2000 *Mathematics Subject Classification* 11F22, 17B69, 17B67, 81T40

A formidable accomplishment last century was determining the explicit list of all finite simple groups. See e.g. [**49**] for more details and references. These groups are:

(i) the cyclic groups C_p, p prime;
(ii) the alternating groups A_n, $n \geq 5$;
(iii) 16 infinite families of groups of Lie type;
(iv) 26 sporadic groups.

An example of a family of finite simple group of Lie type is $\mathrm{PSL}_n(\mathbb{F}_q)$, i.e. the group of $n \times n$ matrices with determinant 1, and entries from the finite field $\mathbb{F}_q$, quotiented out by its centre (the scalar matrices aI, where $a^n = 1$).

The mysteriousness of the sporadics is due to their falling outside those infinite families. They range in size from the Mathieu group M_{11}, with order 7920 and discovered in 1861, to the Monster $\mathbb{M}$, with order

$$|\mathbb{M}| = 2^{45} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \times 10^{53} \ . \tag{2.1}$$

The existence of $\mathbb{M}$ was conjectured in 1973 by Fischer and Griess, and finally constructed in 1980 by Griess [**50**]. Most sporadics arise in $\mathbb{M}$ (e.g. as quotients of subgroups). We'll encounter many of these sporadics in the coming pages, but most of our attention will be directed at $\mathbb{M}$.

Griess showed in fact that $\mathbb{M}$ was the automorphism group of a 196883-dimensional commutative nonassociative algebra, now called the *Griess algebra*, but the construction was somewhat artificial. We now understand [**44**] the Griess algebra as the first nontrivial tier of an infinite-dimensional graded algebra, the *Moonshine module* $V^\natural$, which lies at the heart of Monstrous Moonshine. We'll discuss $V^\natural$ in §4.2; we will find that it has a very rich algebraic structure, is conjectured to obey a strong uniqueness property, and has automorphism group $\mathbb{M}$.

The Monster $\mathbb{M}$ has a remarkably simple presentation. As with any noncyclic finite simple group, it is generated by its involutions (i.e. elements of order 2) and so will be a homomorphic image of a Coxeter group. Let $\mathcal{G}_{pqr}$, $p \geq q \geq r \geq 2$, be the graph consisting of three strands of lengths $p+1, q+1, r+1$, sharing a common endpoint. Label the $p+q+r+1$ nodes as in Figure 1. Given any graph $\mathcal{G}_{pqr}$, define Y_{pqr} to be the group consisting of a generator for each node, obeying the usual Coxeter group relations (i.e. all generators are involutions, and the product gg' of two generators has order 3 or 2, depending on whether or not the two nodes are adjacent), together with one more relation:

$$(ab_1b_2ac_1c_2ad_1d_2)^{10} = 1 \ . \tag{2.2}$$

The groups Y_{pqr}, for $p \leq 5$, have now all been identified (see e.g. [**61**]). Conway conjectured and, building on work by Ivanov [**60**], Norton proved [**98**] that $Y_{555} \cong Y_{444}$ is the 'Bimonster', the wreathed-square $\mathbb{M} \wr C_2$ of the Monster (so has order $2|\mathbb{M}|^2$). A closely related presentation of the Bimonster has 26 involutions as generators and has relations given by the incidence graph of the projective plane of order 3; the Monster itself arises from 21 involutions and the affine plane of order 3 [**25**]. Likewise, $Y_{553} \cong Y_{443} \cong \mathbb{M} \times C_2$. Other sporadics arise in e.g. $Y_{533} \cong Y_{433}$ (the Baby Monster $\mathbb{B}$), $Y_{552} \cong Y_{442}$ (the Fischer group Fi'_{24}), and $Y_{532} \cong Y_{432}$ (the Fischer group Fi_{23}). The Coxeter groups of $\mathcal{G}_{555}$, $\mathcal{G}_{553}$, $\mathcal{G}_{533}$, $\mathcal{G}_{552}$, and $\mathcal{G}_{532}$, are all infinite groups of hyperbolic reflections in e.g. $\mathbb{R}^{17,1}$, and contain copies of groups such as the affine E_8 Weyl group, so the geometry here should be quite pretty. What role, if any, these remarkable presentations have in Moonshine hasn't been established yet. As a first step though, [**97**] identifies in $\mathrm{Aut}(V^\natural)$ the 21 involutions generating $\mathbb{M}$.

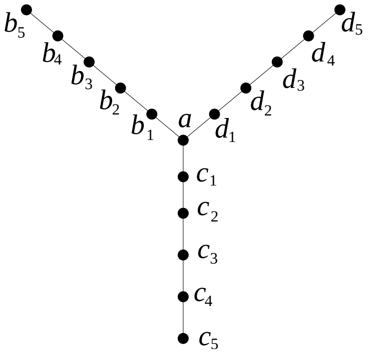


Figure 1. The graph $\mathcal{G}_{555}$ presenting the Bimonster

The Monster has 194 conjugacy classes, and so that number of irreducible representations. Its character table (and much other useful information) is given in the Atlas [**22**], where we also find analogous data for the other simple groups of 'small' order. For example, we find that $\mathbb{M}$ has exactly 2, 3, 4 conjugacy classes of elements of order 2, 3, 4, respectively — these classes are named 2A, 2B, 3A, etc. We also find that the dimensions of the smallest irreducible representations of $\mathbb{M}$ are 1, 196883, 21296876, and 842609326. This is the same 196883 as on the right side of (1.1), and as the dimension of the Griess algebra.

2.2. The j-function. The group $\mathrm{SL}_2(\mathbb{R})$, consisting of 2×2 matrices of determinant 1 with real entries, acts on the upper half-plane $\mathbb{H} := \{\tau \in \mathbb{C} \,|\, \mathrm{Im}(\tau) > 0\}$ by fractional linear transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tau = \frac{a\tau + b}{c\tau + d} \ . \tag{2.3}$$

Of course this is really an action of $\mathrm{PSL}_2(\mathbb{R}) := \mathrm{SL}_2(\mathbb{R})/\{\pm I\}$ on $\mathbb{H}$, but it is more convenient to work with $\mathrm{SL}_2(\mathbb{R})$. $\mathbb{H}$ is the hyperbolic plane, one of the three possible geometries in two dimensions (the others are the sphere and the Euclidean plane), and $\mathrm{PSL}_2(\mathbb{R})$ is its group of orientation-preserving isometries.

Let G be a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$. Then the space $G\backslash\mathbb{H}$ has a natural structure of an orientable surface, and inherits a complex structure from $\mathbb{H}$ (so can be regarded as a complex curve). By the *genus* of the group G, we mean the genus of the resulting real surface $G\backslash\mathbb{H}$. For example, the choice $G = \mathrm{SL}_2(\mathbb{Z})$ yields the sphere with one puncture, so $\mathrm{SL}_2(\mathbb{Z})$ has genus 0. Moreover, any curve Σ with genus g and n punctures, for $3g+n > 3$, is equivalent as a complex curve to the space $G\backslash\mathbb{H}$, for some subgroup G of $\mathrm{SL}_2(\mathbb{R})$ isomorphic to the fundamental group $\pi_1(\Sigma)$.

The most important choice for G is $\mathrm{SL}_2(\mathbb{Z})$, thanks to its interpretation as the modular group of the torus. Most groups G of interest are commensurable with $\mathrm{SL}_2(\mathbb{Z})$, i.e. $G \cap \mathrm{SL}_2(\mathbb{Z})$ has finite index in both G and $\mathrm{SL}_2(\mathbb{Z})$. Examples of these are the congruence subgroups

$$\Gamma(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \mid \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \} \ , \tag{2.4a}$$

$$\Gamma_0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \mid N \text{ divides } c \} \ . \tag{2.4b}$$

For example $\Gamma_0(N)$ has genus 0 for $N = 2, 13, 25$, while $N = 50$ has genus 2 and $N = 24$ has genus 3. The following definition includes all groups arising in Monstrous Moonshine.

Definition 1. *Call a discrete subgroup G of $SL_2(\mathbb{R})$ a* moonshine-type modular group, *if it contains some $\Gamma_0(N)$, and also obeys the condition that*

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in G \text{ iff } t \in \mathbb{Z} \ .$$

Such a modular group is necessarily commensurable with $\mathrm{SL}_2(\mathbb{Z})$. Note that for such a G, any meromorphic function $f : G\backslash\mathbb{H} \to \mathbb{C}$ will have a Fourier expansion of the form $f(\tau) = \sum_{n=-\infty}^{\infty} a_n q^n$, where $q = e^{2\pi \mathrm{i} \tau}$.

Definition 2. *Let G be any subgroup of $SL_2(\mathbb{R})$ commensurable with $SL_2(\mathbb{Z})$. By a* modular function f for G *we mean any meromorphic function $f : \mathbb{H} \to \mathbb{C}$, such that*

$$f(\frac{a\tau + b}{c\tau + d}) = f(\tau) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G \tag{2.5}$$

and such that, for any $A \in \mathrm{SL}_2(\mathbb{Z})$, the function $f(A.\tau)$ has Fourier expansion of the form $\sum_{n=-\infty}^{\infty} b_n q^{n/N}$ for some N and b_n (both depending on A), and where $b_n = 0$ for all but finitely many negative n.

This definition simply states that f is a meromorphic function on the *compact* surface $\Sigma_G := G\backslash\overline{\mathbb{H}}$, where $\overline{\mathbb{H}} := \mathbb{H} \cup \mathbb{Q} \cup \{\mathrm{i}\infty\}$. The G-orbits of $\mathbb{Q} \cup \{\mathrm{i}\infty\}$ are called *cusps*; their role is to fill in the punctures of $G\backslash\mathbb{H}$, compactifying the surface, as there are much fewer meromorphic functions on compact surfaces than on noncompact ones (compare the Riemann sphere to the complex plane!).

We are especially interested in genus 0 groups G of moonshine-type. Their modular functions are particularly easy to characterise: there will be a unique modular function J_G for G, with q-expansion of the form

$$J_G(\tau) = q^{-1} + \sum_{n=1}^{\infty} a_n q^n \ ; \tag{2.6}$$

the modular functions for G are precisely the rational functions $f(\tau) = \frac{\mathrm{poly}(J_G(\tau))}{\mathrm{poly}(J_G(\tau))}$ in J_G. This function J_G is called the (normalised) *Hauptmodul* for the genus 0 group G. For example, the modular group $\mathrm{SL}_2(\mathbb{Z})$ has Hauptmodul

$$J_{\mathrm{SL}_2(\mathbb{Z})}(\tau) = J(\tau) = q^{-1} + 196884\, q + 214\,93760\, q^2 + 8642\,909970\, q^3 + \cdots \ . \tag{2.7}$$

This 196884 is the same as that on the left-side of (1.1). Historically, in place of this Hauptmodul was the equivalent

$$j(\tau)=\frac{(\theta_2(\tau)^8+\theta_3(\tau)^8+\theta_4(\tau)^8)^3}{8\eta(\tau)^{24}}=J(\tau)+744\ .$$

As we know, there are other genus 0 modular groups. For example the Hauptmoduls for $\Gamma_0(2)$, $\Gamma_0(13)$, and $\Gamma_0(25)$, are respectively

$$J_2(\tau)=q^{-1}+276q-2048q^2+11202q^3-49152q^4+184024q^5+\cdots\ , \qquad (2.8a)$$

$$J_{13}(\tau)=q^{-1}-q+2q^2+q^3+2q^4-2q^5-2q^7-2q^8+q^9+\cdots\ , \qquad (2.8b)$$

$$J_{25}(\tau)=q^{-1}-q+q^4+q^6-q^{11}-q^{14}+q^{21}+q^{24}-q^{26}+\cdots\ . \qquad (2.8c)$$

Thompson [**115**] proved there are only finitely many modular groups of moonshine-type in each genus. Cummins [**28**] has found all of these of genus 0 and 1. In particular there are precisely 6486 genus 0 moonshine-type groups. Exactly 616 of these have Hauptmoduls with rational (in fact integral) coefficients, the remainder have cyclotomic integer coefficients. There are some natural equivalences (e.g. a Galois action) which collapse this number to 371, 310 of which have integral Hauptmoduls.

In genus >0, two functions are needed to generate the function field. A complication facing the development of a higher-genus Moonshine is that, unlike the situation in genus 0 considered here, there is no canonical choice for these generators.

See e.g. [**92**] for a very readable account of some of the circle of ideas meandering through this subsection. Modular functions are discussed in e.g. [**75**].

3. The Monstrous Moonshine conjectures

The number on the left of (1.1) is the first nontrivial coefficient of the j-function, and the numbers on the right are the dimensions of the smallest irreducible representations of the Fischer–Griess Monster $\mathbb{M}$. On the one side we have a modular function; on the other, a sporadic finite simple group. Moonshine is the explanation and generalisation of this unlikely connection.

But first, why can't (1.1) merely be a coincidence? This is soon dispelled by comparing the next few coefficients of J with the dimensions of irreducible representations of $\mathbb{M}$:

$$214\,93760=212\,96876+196883+1\ , \qquad (3.1a)$$

$$8642\,99970=8426\,09326+212\,96876+2\cdot 196883+2\cdot 1\ . \qquad (3.1b)$$

3.1. The fundamental conjecture of Conway and Norton. The central structure in the attempt to understand equations (1.1) and (3.1) is an infinite-dimensional graded module for the Monster:

$$V=V_0\oplus V_1\oplus V_2\oplus V_3\oplus\cdots\ . \qquad (3.2a)$$

If we let ρ_d denote the d-dimensional irreducible representation of $\mathbb{M}$, then the first few subspaces will be $V_0=\rho_1$, $V_1=\{0\}$, $V_2=\rho_1\oplus\rho_{196883}$, and $V_3=\rho_1\oplus\rho_{196883}\oplus\rho_{21296876}$. This module is to have graded dimension

$$\dim_V(\tau)=\sum_{n=0} q^n\dim(V_n)=1+196884q^2+214\,93760q^3+\cdots=qJ(\tau)\ . \qquad (3.2b)$$

Of course, (3.2b) alone certainly doesn't uniquely determine V, but assume for now this V has been found. Thompson [**114**] suggested studying in addition the graded traces

$$T_g(\tau) := q^{-1} \sum_{n=0}^{\infty} \mathrm{ch}_{V_n}(g)\, q^n \tag{3.2c}$$

for all $g \in \mathbb{M}$, where the ch_{V_n} are characters. As taking $g = 1$ recovers J, (3.2c) is a natural twist of (3.2b). The functions T_g are now called the *McKay–Thompson series.*

Conway and Norton conjectured [**24**] that for each element g of the Monster $\mathbb{M}$, T_g is the Hauptmodul

$$J_{G_g}(\tau) = q^{-1} + \sum_{n=1}^{\infty} a_n(g)\, q^n \tag{3.3}$$

for a genus 0 subgroup G_g of $\mathrm{SL}_2(\mathbb{R})$. So for each n the coefficient $g \mapsto a_n(g)$ defines a character $\mathrm{ch}_{V_n}(g)$ of $\mathbb{M}$. They explicitly identify each of the groups G_g; these groups each contain $\Gamma_0(N)$ as a normal subgroup, for some N dividing $o(g) \gcd(24, o(g))$ ($o(g)$ is the order of g), and the quotient group $G_g/\Gamma_0(N)$ has exponent 2 (or 1).

Since $T_g = T_{hgh^{-1}}$ by definition, there are at most 194 distinct McKay–Thompson series. All coefficients $a_n(g)$ are integers (as are in fact most entries of the character table of $\mathbb{M}$). This implies that $T_g = T_h$ whenever the cyclic subgroups $\langle g\rangle$ and $\langle h\rangle$ are equal. In fact, the total number of distinct McKay–Thompson series T_g arising in Monstrous Moonshine turns out to be only 171. The first 50 coefficients $a_n(g)$ of each T_g are given in [**91**]. Together with the recursions given in §3.3 below, this allows one to effectively compute arbitrarily many coefficients $a_n(g)$ of the Hauptmoduls. It is also this which uniquely defines V, up to equivalence, as a graded $\mathbb{M}$-module.

For example, there are two different conjugacy classes of order 2 elements. One of these gives the Hauptmodul J_2 in (2.8a), while the other corresponds to (3.4) below. Similarly, (2.8b) corresponds to an order 13 element, but J_{25} in (2.8c) doesn't equal any T_g. Recall that there are exactly 616 Hauptmoduls of moonshine-type with integer coefficients, so most of these don't arise as T_g. Recently [**23**], a fairly simple characterisation has been found of the groups arising as G_g in Monstrous Moonshine. Their proof of this characterisation is by exhaustion.

Conway coined this conjecture *Monstrous Moonshine.* The word 'moonshine' here is English slang for 'insubstantial or unreal', 'idle talk or speculation', 'an illusive shadow'. It is meant to give the impression that matters here are dimly lit, and that [**24**] is 'distilling information illegally' from the character table of $\mathbb{M}$.

Monstrous Moonshine began, unofficially, in 1975 when Andrew Ogg remarked that the list of primes p for which the group

$$\Gamma_0(p)+ := \langle \Gamma_0(p), \frac{1}{\sqrt{p}} \begin{pmatrix} 0 & -1 \\ p & 0 \end{pmatrix} \rangle \tag{3.4}$$

has genus 0, is precisely equal to the list of primes p dividing the order of $\mathbb{M}$. Indeed, in the tables of [**24**] we find that, for each prime p dividing $|\mathbb{M}|$, an element g of $\mathbb{M}$ of order p is assigned the group $G_g = \Gamma_0(p)+$.

3.2. Lie theory and Moonshine. McKay not only noticed (1.1), but also observed that

$$j(\tau)^{\frac{1}{3}} = q^{-\frac{1}{3}} \left(1 + 248q + 4124q^2 + 34\,752q^3 + \cdots\right) . \tag{3.5}$$

The point is that 248 is the dimension of the defining representation of the E_8 Lie group, while $4124 = 3875 + 248 + 1$ and $34\,752 = 30\,380 + 3875 + 2 \cdot 248 + 1$. Incidentally, $j^{\frac{1}{3}}$ is a generating modular function for the genus-0 group $\Gamma(3)$. Thus Moonshine is related somehow to Lie theory.

McKay later found independent relationships with Lie theory [**89**], [**15**], [**47**], reminiscent of his famous A-D-E correspondence with finite subgroups of $\mathrm{SU}_2(\mathbb{C})$. As mentioned earlier, $\mathbb{M}$ has two conjugacy classes of involutions. Let K be the smaller one, called '2A' in [**22**] (the alternative, class '2B', has almost 100 million times more elements). The product of any two elements of K will lie in one of nine conjugacy classes: namely, 1A, 2A, 2B, 3A, 3C, 4A, 4B, 5A, 6A, corresponding respectively to elements of orders 1, 2, 2, 3, 3, 4, 4, 5, 6. It is surprising that, for such a complicated group as $\mathbb{M}$, that list stops at only 6 — we call $\mathbb{M}$ a *6-transposition group* for this reason (more on this in §5.2). The punchline: McKay noticed that those nine numbers are precisely the labels of the affine E_8 diagram (see Figure 2). Thus we can attach a conjugacy class of $\mathbb{M}$ to each vertex of the $\widehat{E}_8$ diagram. An interpretation of the *edges* in the $\widehat{E_8}$ diagram, in terms of $\mathbb{M}$, is unfortunately not known.

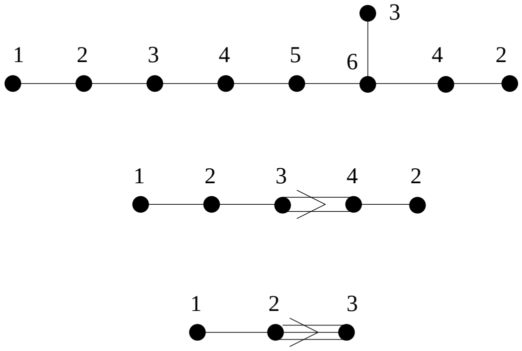


Figure 2. The affine E_8, F_4, and G_2 diagrams with labels

We can't get the affine E_7 labels in a similar way, but McKay noticed that an order two folding of affine E_7 gives the affine F_4 diagram, and we can obtain its labels using the Baby Monster $\mathbb{B}$ (the second largest sporadic). In particular, let K now be the smallest conjugacy class of involutions in $\mathbb{B}$ (also labelled '2A' in [**22**]); the conjugacy classes in KK have orders 1, 2, 2, 3, 4 ($\mathbb{B}$ is a 4-transposition group), and these are the labels of $\widehat{F_4}$. Of course we'd prefer $\widehat{E_7}$ to $\widehat{F_4}$, but perhaps that *two*-folding has something to do with the fact that an order-*two* central extension of $\mathbb{B}$ is the centraliser of an element $g \in \mathbb{M}$ of order *two*.

Now, the *triple*-folding of affine E_6 is affine G_2. The Monster has three conjugacy classes of order *three*. The smallest of these ('3A') has a centraliser which is a *triple* cover of the Fischer group $Fi'_{24}.2$. Taking the smallest conjugacy class of involutions in $Fi'_{24}.2$, and multiplying it by itself, gives conjugacy classes with orders 1, 2, 3 (hence $Fi'_{24}.2$ is a 3-transposition group) — and those not surprisingly are the labels of $\widehat{G_2}$!

Although we now understand (3.5) (see §4.1) and have proven the fundamental Conway–Norton conjecture (see §§4.2–4.4), McKay's $\widehat{E_8}, \widehat{F_4}, \widehat{G_2}$ observations still have no explanation. In [**47**] these patterns are extended, by relating various simple groups to the $\widehat{E_8}$ diagram with deleted nodes.

3.3. Replicable functions. There are several other less important conjectures. One which played an important role in ultimately proving the main conjecture involves the *replication formulae.* Conway–Norton want to think of the Hauptmoduls T_g as being intimately connected with $\mathbb{M}$; if so, then the group structure of $\mathbb{M}$ should somehow directly relate different T_g. Considering the power map $g \mapsto g^n$ leads to the following.

It was well-known classically that $j(\tau)$ has the property that $j(p\tau) + j(\frac{\tau}{p}) + j(\frac{\tau+1}{p}) + \cdots + j(\frac{\tau+p-1}{p})$ is a polynomial in $j(\tau)$, for any prime p (*proof:* it's a modular function for $\mathrm{SL}_2(\mathbb{Z})$, and hence equals a rational function of $j(\tau)$; since its only poles will be at the cusps, the denominator polynomial must be trivial). Hence the same will hold for J. More generally, we get

$$\sum_{ad=n, 0\le b<d} J(\frac{a\tau+b}{d}) = Q_n(J(\tau)) \ , \tag{3.6a}$$

where Q_n is the unique polynomial for which $Q_n(J(\tau)) - q^{-n}$ has a q-expansion with only strictly positive powers of q. For example, $Q_2(x) = x^2 - 2a_1$ and $Q_3(x) = x^3 - 3a_1 x - 3a_2$, where we write $J(\tau) = \sum_n a_n q^n$. The left side of (3.6a) is really a Hecke operator applied to J. These equations (3.6a) can be rewritten into recursions such as $a_4 = a_3 + (a_1^2 - a_1)/2$, or collected together into the remarkable expression (originally due to Zagier)

$$p^{-1} \prod_{\substack{m>0 \\ n\in\mathbb{Z}}} (1 - p^m q^n)^{a_{mn}} = J(z) - J(\tau) \ , \tag{3.6b}$$

where $p = e^{2\pi \mathrm{i} z}$.

Conway and Norton conjectured [**24**] that these formulas have an analogue for any McKay–Thompson series T_g. In particular, (3.6a) becomes

$$\sum_{ad=n, 0\le b<d} T_{g^a}(\frac{a\tau+b}{d}) = Q_{n,g}(T_g(\tau)) \ , \tag{3.7a}$$

where $Q_{n,g}$ plays the same role for T_g that Q_n played for J. These are called the replication formulae. Again, these yield recursions like $a_4(g) = a_2(g) + (a_1(g)^2 - a_1(g^2))/2$, or can be collected into the expression

$$p^{-1} \exp[-\sum_{k>0} \sum_{\substack{m>0 \\ n\in\mathbb{Z}}} a_{mn}(g^k) \frac{p^{mk} q^{nk}}{k}] = T_g(z) - T_g(\tau) \ . \tag{3.7b}$$

This looks a lot more complicated than (3.6b), but you can glimpse the Taylor expansion of $\ln(1 - p^m q^n)$ there and in fact for $g = 1$ (3.7b) reduces to (3.6b).

Axiomatising (3.7a) leads to Norton's notion of *replicable function* [**96**], [**1**]. Write $f^{(1)}(\tau) = q^{-1} + \sum_{k=1}^{\infty} b_k^{(1)} q^k$, and replacing each T_{g^a} in (3.7a) with $f^{(a)}$, use (3.7a) to recursively define each $f^{(n)}$. If each $f^{(n)}$ has a q-expansion of the form $f^{(n)}(\tau) = q^{-1} + \sum_{k=1}^{\infty} b_k^{(n)} q^k$ — i.e. no fractional powers of q arise — then we call $f = f^{(1)}$ replicable. Equation (3.7a) says the McKay–Thompson series are replicable, and [**30**] proved that the Hauptmodul of any genus 0 modular group of moonshine-type is replicable, provided its coefficients are rational. Conversely, Norton conjectured that any replicable function with rational coefficients is either such a Hauptmodul, or one of the 'modular fictions' $f(\tau) = q^{-1}$, $f(\tau) = q^{-1} \pm q$. This conjecture seems difficult and is still open. Incidentally, if the coefficients $b_k^{(1)}$ are irrational, then the definition (3.7a) of replicability should be modified to include Galois automorphisms (see §8 of [**29**]). Replication in positive genus is discussed in [**109**].

Replication (3.7a) concerns the power map $g \mapsto g^n$ in $\mathbb{M}$. Can Moonshine see more of the group structure of $\mathbb{M}$? We explored one step in this direction in §3.2, where McKay modeled products of conjugacy classes using Dynkin diagrams. A different idea is given in §5.1. It would be very desirable to find other direct connections between the group operation in $\mathbb{M}$ and e.g. the McKay–Thompson series.

3.4. The Leech lattice and Moonshine. The Leech lattice $\Lambda = \Lambda_{24}$ is a 24-dimensional even self-dual lattice [**26**] which is to lattices much as the $\mathbb{M}$-module V of (3.2a) turns out to be for vertex operator algebras (see §4.2 below). Λ has no vectors of odd norm, no norm-2 vectors, and precisely 196560 norm-4 vectors — a number remarkably close to the monstrous 196883. In fact its theta series $\Theta_\Lambda(\tau) = \sum_{v \in \Lambda} q^{v \cdot v/2}$, when divided by $\eta(\tau)^{24}$, equals $J(\tau) + 24$. Is this another example of Moonshine?

Indeed it is. However we have:

THEOREM 3. *Let $L \subset \mathbb{R}^n$ be any n-dimensional positive-definite lattice whose norms $\mathbf{v} \cdot \mathbf{v}$ are all rational. Let $\mathbf{t} \in \mathbb{R}^n$ be any vector with finite order in L: i.e. $m\mathbf{t} \in L$ for some nonzero $m \in \mathbb{Z}$. Then the theta series*

$$\Theta_{L+\mathbf{t}}(\tau) := \sum_{\mathbf{v} \in L} e^{\pi \mathrm{i} \tau \, (\mathbf{v}+\mathbf{t})^2} \ ,$$

divided by $\eta(\tau)^n$, is a modular function for some $\Gamma(N)$.

See e.g. Theorem 20 of [**100**] for a proof of this classical result. If L is in fact an *even* lattice (i.e. all norms $\mathbf{v} \cdot \mathbf{v}$ lie in $2\mathbb{Z}$), we can say more. Let $L^* := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot L \subseteq \mathbb{Z}\}$ be the dual lattice. It contains L with finite index; write $\mathbf{t}_i + L$, $i = 1, \ldots, M$, for the finitely many cosets in L^*/L. Define a column vector $\vec{\chi}_L(\tau)$ with ith component $\Theta_{\mathbf{t}_i+L}(\tau)/\eta(\tau)^n$. Then $\vec{\chi}_L$ forms a *vector-valued modular function* for $\mathrm{SL}_2(\mathbb{Z})$: for any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$,

$$\vec{\chi}_L\left(\frac{a\tau + b}{c\tau + d}\right) = \rho(A)\, \vec{\chi}_L(\tau) \tag{3.8}$$

for some M-dimensional unitary matrix representation ρ of $\mathrm{SL}_2(\mathbb{Z})$. In particular, for the Leech lattice $L = \Lambda$, $M = 1$ and we can quickly identify $\Theta_\Lambda(\tau)$ in terms of $J(\tau)$. Although the 196560 ≈ 196884 coincidence is thus trivial to explain, it will turn out to be a very instructive example of Moonshine.

The lattices are related to groups through their automorphism groups, which will always be finite for positive-definite lattices. The automorphism group $Co_0 := \mathrm{Aut}(\Lambda)$ of the Leech lattice has order about 8×10^{18}, and is the direct product of C_2 with Conway's sporadic group Co_1. Several other sporadics are also involved in $\mathrm{Aut}(\Lambda)$. To each automorphism $\alpha \in \mathrm{Aut}(\Lambda)$, let θ_α denote the theta series of the sublattice of Λ fixed by α. [**24**] also associate to each automorphism α a certain function $\eta_\alpha(\tau)$ of the form $\prod_i \eta(a_i\tau)/\prod_j \eta(b_j\tau)$ built out of the Dedekind eta. Both θ_α and η_α are constant on each conjugacy class in $\mathrm{Aut}(\Lambda)$, of which there are 202. [**24**] remarks that the ratio $\theta_\alpha/\eta_\alpha$ always seems to equal some McKay–Thompson series $T_{g(\alpha)}$. See also [**87**].

It turns out that this observation isn't quite correct [**74**]. For each $\alpha \in \mathrm{Aut}(\Lambda)$, the subgroup of $\mathrm{SL}_2(\mathbb{R})$ which fixes $\theta_\alpha/\eta_\alpha$ is indeed genus 0, but for exactly 15 conjugacy classes in $\mathrm{Aut}(\Lambda)$, $\theta_\alpha/\eta_\alpha$ is not the Hauptmodul.

Similarly, one can ask this for the E_8 lattice, whose automorphism group is the Weyl group (of order $\approx 7 \times 10^8$) of the E_8 Lie group. The automorphisms α of the E_8 lattice which yield a Hauptmodul were classified in [**19**].

4. Proof of the Monstrous Moonshine conjectures

At first glance, the significance of the Moonshine conjectures seems very unlikely: they constitute after all a finite set of very specialised coincidences. The whole point though is to try to understand *why* such seemingly incomparable objects as the Monster and the Hauptmoduls can be so related, and to try to extend and apply this understanding to other contexts. Establishing the truth (or falsity) of the conjectures was merely meant as an aid to uncovering the why of Moonshine. Indeed, in achieving this understanding, important new algebraic structures were formulated. We will sketch this theory below.

The main Conway–Norton conjecture was attacked almost immediately. Thompson showed [**113**] (see also [**103**]) that if $g \mapsto a_n(g)$ is a character for all sufficiently small n (apparently $n \leq 1300$ is sufficient), then it will be for all n. He also showed that if certain congruence conditions hold for a certain number of $a_n(g)$ (all with $n \leq 100$), then all $g \mapsto a_n(g)$ will be *virtual* characters (i.e. differences of true characters of $\mathbb{M}$). Atkin, Fong, and Smith (see [**110**] for details) used that to prove on a computer that indeed all $a_n(g)$ were virtual characters (they didn't quite reach $n = 1300$ though). But their work doesn't say anything more about the underlying (possibly virtual) representation V, other than its existence. Their work plays no role in the following.

We want to show that the McKay–Thompson series $T_g(\tau)$ of (3.2c) equals the Hauptmodul $J_{G_g}(\tau)$ in (3.3). First, we need to construct the infinite-dimensional module V of $\mathbb{M}$. We discuss this, and the underlying theory of vertex operator algebras, in §4.2. Borcherds' strategy [**11**] was to bring in Lie theory, by associating to the module V a 'Monster Lie algebra'. This algebra, and the underlying theory of generalised Kac–Moody algebras, is described in §4.3. In the final subsection we go from the Monster Lie algebra to the replication formulae, and conclude the proof. We begin this section though by explaining the much simpler connection of E_8 with $j^{\frac{1}{3}}$.

4.1. E_8 *and* $j^{\frac{1}{3}}$*.* An explanation for the relation between E_8 and $j^{\frac{1}{3}}$ was found almost immediately, by Kac and Lepowsky [**65**], [**76**]: $j^{\frac{1}{3}}$ is the (normalised) character of a representation of the affine Kac–Moody algebra $E_8^{(1)}$. Given a finite-dimensional simple Lie algebra $\mathfrak{g}$, the affine algebra $\mathfrak{g}^{(1)}$ is the infinite-dimensional Lie algebra consisting of all Laurent polynomials $\sum_{n=-\infty}^{\infty} a_n t^n$ where $a_n \in \mathfrak{g}$ and t is an indeterminate, together with a central term and derivation $D = -L_0$ (see [**66**], [**70**]). Highest weight representations are defined in the usual way. Thanks largely to the fact that the affine Weyl group is a semi-direct product of the additive group $\mathbb{Z}^r$ ($r = \mathrm{rank}(\mathfrak{g})$) with the finite Weyl group, the characters of these representations (especially the 'integrable' highest weight ones, which are the direct analogue of the finite-dimensional representations of $\mathfrak{g}$) transform nicely with respect to $\mathrm{SL}_2(\mathbb{Z})$. See e.g. Chapter 13 of [**70**] for details. This is probably the single biggest reason Kac–Moody algebras are so well-known.

THEOREM 4. [**69**] *Let $\mathfrak{g}$ be any finite-dimensional Lie algebra, and $\mathfrak{g}^{(1)}$ denote the corresponding affine algebra. Let P_+^k denote the finitely many 'level k integrable highest weight modules' L_λ of $\mathfrak{g}^{(1)}$. The $\mathfrak{g}^{(1)}$-module L_λ has a natural $\mathbb{Z}$-grading $L_\lambda = \oplus_{n=0}^{\infty} (L_\lambda)_n$ into finite-dimensional $\mathfrak{g}$-modules $(L_\lambda)_n$. Let $\chi_\lambda(\tau) = q^{h_\lambda - c/24} \sum_{n=0}^{\infty} \dim((L_\lambda)_n)\, q^n$ be the corresponding normalised character (for some appropriate choice of $h_\lambda - c/24 \in \mathbb{Q}$). Then each χ_λ is a holomorphic function in $\mathbb{H}$, and the vector $\vec{\chi}_k(\tau)$ with entry $\chi_\lambda(\tau)$ for each $\lambda \in P_+^k$ defines a vector-valued modular function for $SL_2(\mathbb{Z})$, as in (3.8), for some finite-dimensional unitary representation ρ of $SL_2(\mathbb{Z})$.*

In fact each character χ_λ will be a rational function in lattice theta series, and so will be a modular function for some $\Gamma(N)$. It turns out that there is only one level 1 integrable highest-weight representation of $E_8^{(1)}$, and its character equals $j^{\frac{1}{3}}$. The modularity of $j^{\frac{1}{3}}$ is thus predicted by Kac–Moody theory, and the fact that the coefficients are dimensions of E_8 representations is automatic. We will see a simultaneous generalisation of Theorems 3 and 4 next section.

We've already encountered the mysterious normalisation q^{-1} of the McKay–Thompson series, and the $q^{-\frac{1}{3}}$ of the $E_8^{(1)}$ character, and more generally the $q^{h_\lambda - c/24}$ of χ_λ. Many explanations have been provided for this pervasive factor. For example, to [**2**] it is topological in origin, and related to the Atiyah–Singer Index Theorem; a geometric interpretation using determinant line bundles is due to Segal [**107**]. In quantum physics it's called the *conformal anomaly* (a breakdown of manifest conformal symmetry when the classical system is quantised), and is introduced in regularisation as a vacuum energy. Probably the simplest instance of it is the prefactor in the familiar definition $\eta(\tau) = q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n)$ for the Dedekind eta: reading through classical proofs for its modularity we find that '$\frac{1}{24}$' here arises through the combination $\zeta(2)/(2\pi)^2$; η appears also in physics in the partition function of the bosonic string, and that same '$\frac{1}{24}$' arises there via regularisation as $-\zeta(-1)/2$. The equivalence of these two expressions for $\frac{1}{24}$ comes from the functional equation of the Riemann zeta. This same zeta value appears famously in the central term of the Virasoro algebra (4.3), and Bloch [**8**] found other zeta values appearing in other algebras of differential operators, many of which have now been interpreted and generalised (starting with [**77**]) within the vertex operator algebra framework.

Although a direct explanation for Monstrous Moonshine using affine algebras has never been found (and certainly isn't expected), the theory of Kac–Moody algebras influenced every stage of the ultimate proof. For this reason we'll briefly sketch their theory. A simple finite-dimensional Lie algebra $\mathfrak{g}$ is built out of the 3-dimensional algebra sl_2, in a simple way; the Dynkin diagram of $\mathfrak{g}$ encodes the exact presentation. In the identical way, Kac–Moody algebras are also built out of copies of sl_2 — the only difference is that the finite-dimensionality constraint (a positive-definiteness condition on the Cartan matrix) is lifted. Their structure is completely analogous to that of the simple Lie algebras: e.g. it has a grading by roots into finite-dimensional spaces; it has a triangular decomposition (making Verma modules possible); it has an invariant symmetric bilinear form. See e.g. [**66**], [**70**] for details. The affine algebras are the class of Kac–Moody algebras especially analogous to the finite-dimensional ones.

4.2. The Moonshine module $V^\natural$. A vital component of the Monstrous Moonshine conjectures came a few years after [**24**]. In a deep work, Frenkel–Lepowsky–Meurman [**43**], [**44**] constructed a graded infinite-dimensional representation $V^\natural$ of $\mathbb{M}$ and conjectured (correctly) that it is the representation V in (3.2a). $V^\natural$ has a very rich algebraic structure: it is in fact a *vertex operator algebra!*

A vertex operator algebra [**9**], [**44**], [**67**], [**39**], [**79**] is an infinite-dimensional vector space V with infinitely many heavily constrained bilinear products $u *_n v$. The name means 'algebra of (generalised) vertex operators'; vertex operators are formal differential operators which originally appeared in physics as quantum fields describing the creation and propagation of physical strings (see §6 below), and were constructed later but independently by Lie theorists (starting with Lepowsky and Wilson) to realise affine Kac–Moody algebras as algebras of differential operators. Because there were vertex operator constructions associated to lattices, affine algebra modules, and string theory, and all of these have connections to modular functions, it was natural to use vertex operators to try to construct the $\mathbb{M}$-module V of (3.2a).

The definition of vertex operator algebra (VOA) is too complicated to give in detail here. A VOA is a graded infinite-dimensional vector space $\mathcal{V} = \oplus_{n=0}^{\infty}\mathcal{V}_n$, where each $\mathcal{V}_n$ is finite-dimensional. To simplify the discussion, we will limit ourselves in this paper to VOAs with one-dimensional $\mathcal{V}_0$, which is typical of the examples relevant to Moonshine (and conformal field theory). To each vector $v \in \mathcal{V}$ we assign a *vertex operator* $Y(v,z)$, which is a formal power series $Y(v,z) = \sum_{m\in\mathbb{Z}} v_{(m)} z^{-m-1}$, with coefficients $v_{(m)} \in \mathrm{End}(\mathcal{V})$. The vertex operator is just the generating function for the products: $u *_n v = u_{(n)}(v)$. These products respect the grading — in particular,

$$\mathcal{V}_k *_n \mathcal{V}_\ell \subseteq \mathcal{V}_{k+\ell-n-1} \ . \tag{4.1}$$

A key axiom, which collects together all the identities obeyed by the products $u *_n v$, can be written as

$$(z-w)^M [Y(u,z), Y(v,w)] = 0 \qquad \forall u, v \in \mathcal{V} \ , \tag{4.2a}$$

for some integer M (depending on u, v), where the bracket in (4.2a) means the commutator $Y(u,z)\,Y(v,w) - Y(v,w)\,Y(u,z)$. This strange-looking formula really says that each such commutator is a linear combination of Dirac deltas and their derivatives, all centred at $z = w$ (see e.g. Corollary 2.2 in [**67**]). Equation (4.2a) implies more down-to-earth identities, such as

$$(u_\ell v)_n = \sum_{i\geq 0}(-1)^i \binom{\ell}{i} \left(u_{\ell-i}\circ v_{n+i} - (-1)^\ell v_{\ell+n-i}\circ u_i\right) \ . \tag{4.2b}$$

There are two distinguished elements in $\mathcal{V}$: the identity $\mathbf{1}\in\mathcal{V}_0$ (so $\mathcal{V}_0=\mathbb{C}\mathbf{1}$) and the *conformal vector* $\omega\in\mathcal{V}_2$. The identity obeys $Y(\mathbf{1},z)=id$, i.e. $\mathbf{1}_{(n)}v=\delta_{n,-1}v$. More interesting is the conformal vector: writing $L_n=\omega_{(n+1)}$, the operators L_n are required to form a representation on $\mathcal{V}$ of the *Virasoro algebra*:

$$[L_m,L_n]=(m-n)L_{m+n}+\delta_{m,-n}\frac{m^3-m}{12}c\,id_{\mathcal{V}} \ , \tag{4.3}$$

for some number $c\in\mathbb{R}$ (an important numerical invariant of $\mathcal{V}$) called the *rank* or *central charge* of the VOA. In addition we require $L_0u=nu$ whenever $u\in\mathcal{V}_n$, and L_{-1} acts on $\mathcal{V}$ as a derivation.

The appearance here of the Virasoro algebra is fundamental. It is the unique nontrivial central extension of the (polynomial) vector fields on S^1, which in turn is the Lie algebra associated to the group Diff(S^1) of diffeomorphisms of the circle.

The notion of a VOA may seem very arbitrary, but as we'll mention in §6 it is the 'chiral algebra' of a conformal field theory. The simplest (and least interesting) special case of a VOA occurs when $M=0$ in (4.2a) — i.e. when all vertex operators commute. Then it is not hard to see that, for each choice of $z\neq 0$, $\mathcal{V}$ would be a commutative associative algebra with unit, whose product is given by $u *_z v := Y(u,z)v$. A more honest way to motivate VOAs has been suggested by Huang: binary trees can be used to keep track of the brackets in nested products, e.g. $a((bc)d)$, and e.g. Lie algebras can be easily formulated using this language [**58**]; in a monumental work [**59**], Huang 'two-dimensionalised' this Lie algebra formulation by replacing binary trees with spheres with tubes, and showed that the result is equivalent to a VOA.

A relation between VOAs and Lie algebras also exists at a more elementary level. Placing $\ell=n=0$ in (4.2b), and evaluating it on the right by $w\in\mathcal{V}$, gives

$$(u_0v)_0w=u_0(v_0w)-v_0(u_0w) \ . \tag{4.4}$$

Writing $[xy]$ for x_0y, this becomes the Lie algebra Jacobi identity, at least when $[xy]=-[yx]$. There are different ways to obtain from this a true Lie algebra. The simplest is that V_1 will be a Lie algebra for that bracket; moreover, the number $\langle u,v\rangle$ defined by $u_1v=\langle u,v\rangle\mathbf{1}$ is an invariant bilinear form for this Lie algebra (by (4.2b) with $\ell=0,n=1$). Equation (4.4) tells us each $\mathcal{V}_n$ is a $\mathcal{V}_1$-module, and in fact e^u will be an automorphism of $\mathcal{V}$ for any $u\in\mathcal{V}_1$. In the most common examples, $\mathcal{V}_1$ will be reductive (i.e. a direct sum of simple and abelian Lie algebras).

Modules M of $\mathcal{V}$ can be defined in the obvious way [**42**], [**79**] — e.g. for each $u\in\mathcal{V}$, each $u_{(n)}$ will be in End(M). All (irreducible) modules M come with a $\mathbb{Z}$-grading $M=\oplus_{n=0}^\infty M_n$, where $u_kM_n\subseteq M_{n+\ell-k-1}$ for all $u\in\mathcal{V}_\ell$, and $L_0x=(n+h)x$ for any $x\in M_n$, where h is some number (the *conformal weight*) depending only on M. The (normalised) character of M is

$$\chi_M(\tau)=q^{-c/24}\mathrm{Tr}_Mq^{L_0}=q^{h-c/24}\sum_{n=0}^\infty \dim(M_n)\,q^n \ .$$

It takes a little effort to construct even the simplest examples of VOAs. The best-behaved ones are called *rational VOAs* [**124**] (borrowing on terminology from physics) and have only finitely many irreducible modules. A rational VOA is associated to any even positive-definite lattice L, and their modules are in one-to-one correspondence with the cosets L^*/L. Another important example: to any affine nontwisted Kac–Moody algebra and choice of positive integer k (the 'level'), the highest weight module $L_{k\Lambda_0}$ has a natural VOA structure, and its modules are precisely the affine algebra modules L_λ for each highest weight $\lambda \in P_+^k$.

One of the deepest results in the theory of VOAs is due to Zhu:

Theorem 5. [**124**] *Let $\mathcal{V}$ be a rational VOA. Its characters $\chi_M(\tau)$ are holomorphic in $\mathbb{H}$, and the subspaces M_n carry representations for Aut($\mathcal{V}$). Write $\vec{\chi}_{\mathcal{V}}(\tau)$ for the vector whose components are the characters $\chi_M(\tau)$ of irreducible modules M. Then $\vec{\chi}_{\mathcal{V}}$ is a vector-valued modular function for $SL_2(\mathbb{Z})$.*

It is believed that the characters $\chi_M(\tau)$ themselves will be modular functions for some $\Gamma(N)$; significant progress towards this was made in [**5**] (see also [**71**]). The proof of Zhu's Theorem is much more difficult than that of Theorem 4, which it generalises.

The automorphism group $\mathrm{Aut}(\mathcal{V})$ is by definition required to fix ω, which is why it respects the grading of $\mathcal{V}$. $\mathrm{Aut}(\mathcal{V})$ is how group theory impinges on VOA theory. Since the automorphism group $\mathrm{Aut}(\mathcal{V})$ of a VOA contains $e^{\mathcal{V}_1}$ as a (normal) subgroup, $\mathrm{Aut}(\mathcal{V})$ can be finite only when $\mathcal{V}_1 = 0$. Zhu's Theorem tells us that Moonshine (without the genus-0 aspect) will hold between the group $\mathrm{Aut}(\mathcal{V})$ and the functions $\chi_M(\tau)$, for any rational VOA.

The most famous example of a VOA is the Moonshine module $V^\natural$ of [**44**]. It is the orbifold of the Leech lattice VOA $\mathcal{V}_\Lambda$ by the ±1-symmetry of Λ, which means it's the direct sum of two parts: an invariant part $V_+^\natural$ and a twisted part $V_-^\natural$ (more on this in §5.1). The orbifold serves two purposes: it removes the constant term '24' from the graded dimension $J + 24$ (hence the subspace $(\mathcal{V}_\Lambda)_1$) of $\mathcal{V}_\Lambda$; and it enhances the symmetry from the discrete part of $\mathrm{Aut}(\mathcal{V}_\Lambda)$, which is an extension of Co_0 by $(C_2)^{24}$, to all of $\mathbb{M}$.

A major claim of [**44**] was that $V^\natural$ is a 'natural' structure (hence their notation). Even so, this bipartite structure to $V^\natural$ complicates its study. We have $V_0^\natural = \mathbb{C}\mathbf{1}$, as usual, but the Lie algebra $\mathcal{V}_1 = \{0\}$ is trivial. For any such VOA, the space $\mathcal{V}_2$ will be a commutative nonassociative algebra with product $u \times v := u_1 v$ and identity $\frac{1}{2}\omega$. For the Moonshine VOA $V^\natural$, this can be shown (with effort!) to be the 196883-dimensional Griess algebra extended by an identity element. From this, we find the automorphism group of $V^\natural$ to be the Monster $\mathbb{M}$. The only irreducible module for $V^\natural$ is itself — such a VOA is called *holomorphic*. Together with Zhu's Theorem, this implies that its character, namely $J(\tau)$, must be a modular function for $\mathrm{SL}_2(\mathbb{Z})$ (strictly speaking, we only get invariance up to a 1-dimensional character of $\mathrm{SL}_2(\mathbb{Z})$, but it is easy to show that character must be identically 1). We'll see in §5.1 how to obtain the other McKay–Thompson series from $V^\natural$.

Conjecturally, there are 71 holomorphic VOAs with rank $c = 24$ [**106**]. Much as the Leech lattice is the unique even self-dual positive-definite lattice of dimension 24 containing no norm-2 vectors [**26**], the Moonshine module $V^\natural$ is *conjecturally* [**44**] the unique holomorphic VOA with $c = 24$ and with trivial $\mathcal{V}_1$. Thus, just as the Leech lattice is the unique lattice with theta series Θ_Λ, so (conjecturally) is the Moonshine module the unique holomorphic VOA with (normalised) graded dimension J. Proving this is one of the most important (and difficult) challenges in the subject.

4.3. The Monster Lie algebra $\mathfrak{m}$. To show that all of the McKay–Thompson series T_g are indeed Hauptmoduls, Borcherds needed identities satisfied by their q-expansions. He obtained these through a Lie algebra he associated to $V^\natural$. Before discussing it, let's briefly describe Borcherds' generalisation of Kac–Moody algebras [**10**].

A Borcherds–Kac–Moody algebra differs from a Kac–Moody algebra in that it is built up from Heisenberg algebras as well as sl_2, and these subalgebras intertwine in more complicated ways. Nevertheless much of the theory for finite-dimensional simple Lie algebras continues to find an analogue in this much more general setting (e.g. root-space decomposition, Weyl group, character formula,...). This unexpected fact is the point of Borcherds–Kac–Moody algebras. For reasons of space we avoid giving here the fairly simple definition, but for this and much more see the review articles [**53**], [**63**], [**102**].

Their basic structure theorem is that of Kac–Moody algebras. In particular, there is a grading by roots into finite-dimensional spaces (except that the 0-graded piece, corresponding to the Cartan subalgebra, may be infinite-dimensional). They also have a triangularisable decomposition and an invariant symmetric bilinear form. Indeed, these structural properties *characterise* Borcherds–Kac–Moody algebras. In this sense Borcherds–Kac–Moody algebras are the ultimate generalisation of simple Lie algebras, in that any further generalisation would lose some basic structural ingredient.

In short, Borcherds' algebras strongly resemble the Kac–Moody ones and constitute a natural and nontrivial generalization. The main differences are that they can be generated by copies of the 3-dimensional Heisenberg algebra as well as sl_2, and that there can be imaginary simple roots. Borcherds introduced these algebras and developed their theory in order to understand the Monster Lie algebra $\mathfrak{m}$.

We want to construct $\mathfrak{m}$ from the Moonshine module $V^\natural = V_0^\natural \oplus V_1^\natural \oplus \cdots$. For later convenience, relabel its subspaces $V^i := V_{i+1}^\natural$. Of course the obvious choice $V_1^\natural = V^0$ is 0-dimensional, so we must modify $V^\natural$ first. Let $II_{1,1}$ denote the even self-dual indefinite lattice consisting of all pairs $(m,n) \in \mathbb{Z}^2$ with inner product $(m,n)\cdot(m',n') = mn' + nm'$. Because it is indefinite, the usual construction of a VOA from a lattice will fail here to produce a true VOA, but most properties will be obtained. Call this near-VOA, $\mathcal{V}_{1,1}$.

The Monster Lie algebra $\mathfrak{m}$ is a Lie algebra associated to the near-VOA $V^\natural \otimes \mathcal{V}_{1,1}$ — see [**11**] for the details. $\mathfrak{m}$ inherits a $II_{1,1}$-grading from $\mathcal{V}_{1,1}$, and this is its root space decomposition: the (m,n) root space is isomorphic (as a vector space) to V^{mn}, if $(m,n) \neq (0,0)$; the $(0,0)$ piece is isomorphic to $\mathbb{R}^2$. Structurally, the Monster Lie algebra has a decomposition $\mathfrak{m} = u^+ \oplus \mathrm{gl}_2 \oplus u^-$ into a sum of Lie subalgebras, where $u^\pm$ are free Lie algebras (see e.g. [**64**]). It inherits the action of $\mathbb{M}$ from $V^\natural$.

This construction of $\mathfrak{m}$ may seem indirect; an alternate approach, anticipated in [**11**] and [**12**], uses *Moonshine cohomology* [**81**] — a functor, inspired by BRST cohomology in conformal field theory, assigning to certain $c = 2$ near-VOAs some Lie algebra carrying an action of $\mathbb{M}$. To $\mathcal{V}_{1,1}$ this functor associates $\mathfrak{m}$.

4.4. Denominator identities and modular equations. It was discovered early on that the Hauptmoduls all obey the replication formulae, and that anything obeying those formulae will be determined by their first few coefficients. The idea then is to show that the McKay–Thompson series T_g of (3.2c) also are replicable. Borcherds did this using Lie algebra denominator identities [**11**].

Finite-dimensional simple Lie algebras $\mathfrak{g}$ possess a very useful formula for their characters, due to Weyl: the (formal) character χ_λ of a module L_λ equals

$$\chi_\lambda := \sum_\mu \dim(L_\lambda(\mu))\, e^\mu = e^{-\rho}\,\frac{\sum_{w\in W}\epsilon(w)\, e^{w(\lambda+\rho)}}{\prod_{\alpha\in\Delta_+}(1-e^{-\alpha})}\ , \tag{4.5}$$

where W is the Weyl group, Δ_+ the positive roots, $\epsilon(w) = \det(w)$ is a sign, and where $\oplus_\mu L_\lambda(\mu)$ is the weight-space decomposition of L_λ. As the weights μ by definition lie in the dual $\mathfrak{h}^*$ of the Cartan subalgebra of $\mathfrak{g}$, the character χ_λ can be regarded as a complex-valued function on the space $\mathfrak{h} \cong \mathbb{C}^r$ ($r = \mathrm{rank}(\mathfrak{g})$).

Consider the trivial representation: i.e. $x \mapsto 0$ for all $x \in \mathfrak{g}$. Its character χ_0 will be identically 1. Thus the character formula (4.5) tells us that a certain alternating sum over a Weyl group, equals a certain product over positive roots. These formulas, called *denominator identities*, are nontrivial even in this finite-dimensional case.

In a famous paper [**83**], Macdonald generalised the denominator identity for (4.5), to infinite sum/product identities, corresponding to the extended Dynkin diagrams. The simplest one was known classically as the Jacobi triple product identity:

$$\sum_{n=-\infty}^{\infty}(-1)^n x^{n^2} y^n = \prod_{m=1}^{\infty}(1-x^{2m})(1-x^{2m-1}y)(1-x^{2m-1}y^{-1})\ . \tag{4.6}$$

Macdonald's identities were later reinterpreted, by Kac and Moody, as denominator identities for the affine algebras. For example, we now know (4.6) to be the denominator identity for the algebra $A_1^{(1)}$.

In particular, the same formula (4.5) holds for Kac–Moody algebras, except that the sum and product are now infinite, the positive roots now come with multiplicities, and the characters are usually normalised by a prefactor $q^{h_\lambda - c/24}$. The variable τ in Theorem 4 is one of the coordinates in the Cartan subalgebra $\mathbb{C}^{r+2}$ of the affine algebra (see e.g. equation (13.2.4) of [**66**]). In that theorem we dropped the remaining variable dependence of the χ_λ for readability, although those additional coordinates serve the important role of guaranteeing linear independence of the characters, and of giving us an action of $\mathrm{SL}_2(\mathbb{Z})$ rather than merely $\mathrm{PSL}_2(\mathbb{Z})$.

Because a Borcherds–Kac–Moody algebra $\mathfrak{g}$ is triangularisable, highest weight $\mathfrak{g}$-modules can be defined in the usual way from Verma modules. The character formula becomes

$$\chi_\lambda = e^{-\rho}\,\frac{\sum_{w\in W}\epsilon(w)\, w(e^{\lambda+\rho}S_\lambda)}{\prod_{\alpha\in\Delta_+}(1-e^{-\alpha})^{\mathrm{mult}\,\alpha}}\ , \tag{4.7}$$

where S_λ is a correction factor due to imaginary simple roots.

The corresponding denominator identity of the Monster Lie algebra $\mathfrak{m}$ can be computed, and is given in (3.6b). Its Weyl group is C_2 and sends the (m, n)-root space to

(n,m); the (m,n) root has multiplicity given by coefficient a_{mn} of J; for each $n>0$ we have an imaginary simple root $(1,n)$ with multiplicity a_n. Because of a cohomological interpretation of all denominator identities, (3.6b) can be 'twisted' by each $g\in\mathbb{M}$, and this gives (3.7b). These formulas are equivalent to the replication formula (3.7a) conjectured in §3.3.

Identities equivalent to (3.7b) were obtained by more elementary means — i.e. methods requiring less of the theory of Borcherds–Kac–Moody algebras — in [**64**] and [**68**], permitting a simplification of Borcherds' proof at this stage.

Now, it turns out that if we verify for each conjugacy class K_g of $\mathbb{M}$ that the first, second, third, fourth and sixth coefficients of the McKay–Thompson series T_g and the corresponding Hauptmodul J_{G_g} agree, then indeed $T_g=J_{G_g}$. That is precisely what Borcherds then did: he compared finitely many coefficients, and as they all equalled what they should, this concluded the proof [**11**] of Monstrous Moonshine!

However, this case-by-case verification occurred at the critical point where the McKay–Thompson series were being compared directly to the Hauptmoduls, and so provides little insight into why the T_g are genus 0. Fortunately a more conceptual explanation of their equality has since been found.

A function f obeying the replication formulae (3.7a) will also obey *modular equations* — i.e. a 2-variable polynomial identity satisfied by $f(x)$ and $f(nx)$. The simplest examples come from the exponential and cosine functions: note that for any $n>0$, $\exp(nx)=(\exp(x))^n$ and $\cos(nx)=T_n(\cos(x))$ where T_n is a Tchebychev polynomial. It was known classically that j (hence J) satisfied a modular equation for any n: e.g. put $X=J(\tau)$ and $Y=J(2\tau)$, then

$$\begin{aligned}(X^2-Y)(Y^2-X)=&\,393768\,(X^2+Y^2)+42987520\,XY+40491318744\,(X+Y)\\&-120981708338256\ .\end{aligned}$$

The only functions $f(\tau)=q^{-1}+a_1q+\cdots$ which obey modular equations for all n, are $J(\tau)$ and the 'modular fictions' q^{-1} and $q^{-1}\pm q$ (which are essentially exp, cos, and sin) [**72**]. More generally, we have:

THEOREM 6. [**29**] *A function $B(\tau)=q^{-1}+\sum_{n=1}^{\infty}b_nq^n$ which obeys a modular equation for all $n\equiv 1$ (mod N), will either be of the form $B(\tau)=q^{-1}+b_1q$, or will be a Hauptmodul for a modular group of moonshine-type.*

The converse is also true [**29**]. The denominator identity argument tells us each T_g obeys a modular equation for each $n\equiv 1$ modulo the order of g, so Theorem 6 then concludes the proof of Monstrous Moonshine.

The computer searches in [**20**] suggest that the hypothesis of Theorem 6 may be considerably weakened, perhaps all the way down to the existence of modular equations for any two distinct primes.

5. Further developments

5.1. Orbifolds. About a third of the McKay–Thompson series T_g will have some negative coefficients. In §5.4 we'll see Borcherds interpret them as dimensions of superspaces (which come with signs). In an important announcement [**97**], on par with [**24**], Norton proposed that, although $T_g(-1/\tau)$ will not usually be another McKay–Thompson series, it will always have nonnegative integer q-coefficients, and these can be interpreted as ordinary dimensions. In the process, he extended the $g \mapsto T_g$ assignment to commuting pairs $(g,h) \in \mathbb{M} \times \mathbb{M}$.

In particular, to each such pair we have a function $N(g,h;\tau)$, which we will call a *Norton series*, such that

$$N(g^a h^c, g^b h^d; \tau) = \alpha\, N(g,h;\frac{a\tau+b}{c\tau+d}) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ , \tag{5.1}$$

for some root of unity α (of order dividing 24, and depending on g, h, a, b, c, d). The Norton series $N(g,h;\tau)$ is either constant, or generates the modular functions for a genus-0 subgroup of $\mathrm{SL}_2(\mathbb{Z})$ containing some $\Gamma(N)$ (but otherwise not necessarily of moonshine-type). Constant $N(g,h;\tau)$ arise when all elements of the form $g^a h^b$ (for $\gcd(a,b) = 1$) are 'non-Fricke' (an element $g \in \mathbb{M}$ is called *Fricke* if the group G_g contains an element sending 0 to $\mathrm{i}\infty$ — the identity 1 is Fricke, as are 120 of the 171 G_g). Each $N(g,h;\tau)$ has a $q^{\frac{1}{N}}$-expansion for that N; the coefficients of this expansion are characters evaluated at h of some central extension of the centralizer $C_{\mathbb{M}}(g)$. Simultaneous conjugation of g, h leaves the Norton series unchanged: $N(aga^{-1}, aha^{-1};\tau) = N(g,h;\tau)$.

For example, when $\langle g,h\rangle \cong C_2 \times C_2$ and g, h, gh are all in class 2A, then $N(g,h;\tau) = \sqrt{J(\tau) - 984}$. The McKay–Thompson series are recovered by the $g = 1$ specialisation: $N(1,h;\tau) = T_h(\tau)$. This action (5.1) of $\mathrm{SL}_2(\mathbb{Z})$ is related to its natural action on the fundamental group $\mathbb{Z}^2$ of the torus, as we'll see in §6, as well as a natural action of the braid group, as we'll see next subsection. Norton arrived at his conjecture empirically, by studying the data of Queen (see §5.3).

The basic tool we have for approaching Moonshine conjectures is the theory of VOAs, so we need to understand Norton's suggestion from that point of view. For reasons of space, we'll limit this discussion to $V^\natural$, but it generalises. Given any automorphism $g \in \mathrm{Aut}(V^\natural)$, we can define g-twisted modules in the obvious way [**36**]. Then for each $g \in \mathbb{M}$, there is a unique g-twisted module, call it $V^\natural(g)$, for $V^\natural$ — this statement generalises the holomorphicity of $V^\natural$ mentioned in §4.2. More generally, given any automorphism $h \in \mathrm{Aut}(V^\natural)$ commuting with g, h will yield an automorphism of $V^\natural(g)$, so we can perform Thompson's twist (3.2c) and write

$$q^{-c/24}\mathrm{Tr}_{V^\natural(g)} h\, q^{L_0} =: \mathcal{Z}(g,h;\tau) \ . \tag{5.2}$$

These $\mathcal{Z}(g,h)$'s can be thought of as the building blocks of the graded dimensions of various eigenspaces in $V^\natural(g)$: e.g. if h has order m, then the subspace of $V^\natural(g)$ fixed by automorphism h will have graded dimension $m^{-1}\sum_{i=1}^m \mathcal{Z}(g,h^i)$. In the case of the Monster considered here, we have $\mathcal{Z}(g,h) = N(g,h)$.

The important paper [**36**] proves that, whenever the subgroup $\langle g,h\rangle$ generated by g and h is cyclic, then $N(g,h)$ will be a Hauptmodul satisfying (5.1). One way this will happen of course is whenever the orders of g and h are coprime. Extending [**36**] to all commuting pairs g, h is one of the most pressing tasks in Moonshine.

This orbifold construction is the same as was used to construct $V^\natural$ from $\mathcal{V}_\Lambda$: $V^\natural$ is the sum of the 'ι'-invariant subspace $V^\natural_+$ of $\mathcal{V}_\Lambda$ with the 'ι'-invariant subspace $V^\natural_-$ of the unique '-1'-twisted module for $\mathcal{V}_\Lambda$, where $\iota \in \mathrm{Aut}(\Lambda)$ is some involution. The graded dimensions of $V^\natural_\pm$ are $2^{-1}(\mathcal{Z}(\pm 1, 1) + \mathcal{Z}(\pm 1, \iota))$, respectively, and these sum to J.

The orbifold construction is also involved in an interesting reformulation of the Hauptmodul property, due to Tuite [**116**]. Assume the uniqueness conjecture: $V^\natural$ is the only VOA with graded dimension J. He argues from this that, for each $g \in \mathbb{M}$, T_g will be a Hauptmodul iff the only orbifolds of $V^\natural$ are $\mathcal{V}_\Lambda$ and $V^\natural$ itself. In e.g. [**62**], this analysis is extended to some of Norton's $N(g,h)$'s, where the subgroup $\langle g, h\rangle$ is not cyclic (thus going beyond [**36**]), although again assuming the uniqueness conjecture.

5.2. Why the Monster? That $\mathbb{M}$ is associated with *modular functions* can be explained by it being the automorphism group of the Moonshine VOA $V^\natural$. But what is so special about this group $\mathbb{M}$ that these modular functions T_g and $N(g,h)$ should be Hauptmoduls? This is still open. One approach is due to Norton, and was first (rather cryptically) stated in [**97**]: the Monster is probably the largest (in a sense) group with the 6-transposition property. Recall from §3.2 that a k-transposition group G is one generated by a conjugacy class K of involutions, where the product gh of any two elements of K has order $\leq k$. For example, taking K to be the transpositions in the symmetric group $G = S_n$, we find that S_n is 3-transposition.

A transitive action of $\Gamma := \mathrm{PSL}_2(\mathbb{Z})$ on a finite set X with one distinguished point $x_0 \in X$, is equivalent to specifying a finite index subgroup Γ_0 of Γ. In particular, Γ_0 is the stabiliser $\{g \in \Gamma \mid g.x_0 = x_0\}$ of x_0, X can be identified with the cosets $\Gamma_0\backslash\Gamma$, and x_0 with the coset Γ_0. (If we avoid specifying x_0, then Γ_0 will be identified only up to conjugation.)

To such an action, we can associate an interesting triangulation of the closed surface $\Gamma_0\backslash\overline{\mathbb{H}}$, called a (modular) *quilt*. The definition, originally due to Norton and further developed by Parker, Conway, and Hsu, is somewhat involved and will be avoided here (but see especially Chapter 3 of [**57**]). It is so-named because there is a polygonal 'patch' covering every cusp of $\Gamma_0\backslash\mathbb{H}$, and the closed surface is formed by sewing together the patches along their edges ('seams'). There are a total of $2n$ triangles and n seams in the triangulation, where n is the index $\|\Gamma_0\backslash\Gamma\| = \|X\|$. The boundary of each patch has an even number of edges, namely the double of the corresponding cusp width. The familiar formula

$$\gamma = \frac{n}{12} - \frac{n_2}{4} - \frac{n_3}{3} - \frac{n_\infty}{2} + 1$$

for the genus γ of $\Gamma_0\backslash\mathbb{H}$ in terms of the index n and the numbers n_i of Γ_0-orbits of fixed-points of order i, can be interpreted in terms of the data of the quilt (see (6.2.3) of [**57**]), and we find in particular that if every patch of the quilt has at most 6 sides, then the genus will be 0 or 1, and genus 1 only exceptionally.

In particular, we're interested in one class of these Γ-actions (actually an $\mathrm{SL}_2(\mathbb{Z})$-action, but this doesn't matter). Recall that the braid group B_3 has presentation

$$\langle \sigma_1, \sigma_2 \mid \sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2\rangle \ , \tag{5.3a}$$

and centre $Z = \langle(\sigma_1\sigma_2\sigma_1)^2\rangle$ [**7**]. It is related to the modular group by

$$B_3/Z \cong \mathrm{PSL}_2(\mathbb{Z}) \ , \quad B_3/\langle(\sigma_1\sigma_2\sigma_1)^4\rangle \cong \mathrm{SL}_2(\mathbb{Z}) \ . \tag{5.3b}$$

Fix a finite group G (we're most interested in the choice $G = \mathbb{M}$). We can define a right action of B_3 on triples $(g_1, g_2, g_3) \in G^3$ by

$$(g_1, g_2, g_3)\sigma_1 = (g_1 g_2 g_1^{-1}, g_1, g_3) \ , \qquad (g_1, g_2, g_3)\sigma_2 = (g_1, g_2 g_3 g_2^{-1}, g_2) \ . \tag{5.4a}$$

We will be interested in this action on the subset of G^3 where all $g_i \in G$ are involutions. The action (5.4a) is equivalent to a reduced version, where we replace (g_1, g_2, g_3) with $(g_1 g_2, g_2 g_3) \in G^2$. Then (5.4a) becomes

$$(g, h)\sigma_1 = (g, gh) \ , \qquad (g, h)\sigma_2 = (gh^{-1}, h) \ . \tag{5.4b}$$

These B_3 actions come from specialisations of the Burau and reduced Burau representations [**7**], respectively, and generalise to actions of B_n on G^n and G^{n-1}. We can get an action of $\mathrm{SL}_2(\mathbb{Z})$ from the B_3 action (5.4b) in two ways: either

(i) by restricting to commuting pairs g, h; or

(ii) by identifying each pair (g, h) with all its conjugates (aga^{-1}, aha^{-1}).

Norton's $\mathrm{SL}_2(\mathbb{Z})$ action of §5.1 arises from the B_3 action (5.4b), when we perform both (i) and (ii).

The quilt picture was designed for this $\mathrm{SL}_2(\mathbb{Z})$ action. The point of this construction is that the number of sides in each patch is determined by the orders of the corresponding elements g, h. If G is say a 6-transposition group (such as the Monster), and we take the involutions g_i from 2A, then each patch will have ≤ 6 sides, and the corresponding genus will be 0 (usually) or 1 (exceptionally if at all). In this way we can relate the Monster with a genus-0 property.

Based on the actions (5.4), Norton anticipates some analogue of Moonshine valid for noncommuting pairs. CFT considerations ('higher genus orbifolds') alluded to in §6 suggest that more natural should be e.g. quadruples $(g, g', h, h') \in \mathbb{M}^4$ obeying $ghg^{-1}h^{-1} = h'g'h'^{-1}g'^{-1}$.

An interesting question is, how much does Monstrous Moonshine determine the Monster? How much of $\mathbb{M}$'s structure can be deduced from e.g. McKay's $\widehat{E}_8$ Dynkin diagram observation, and/or the (complete) replicability of the T_g, and/or Norton's conjectures in §5.1, and/or Modular Moonshine in §5.4 below? A small start toward this is taken in [**99**], where some control on the subgroups of $\mathbb{M}$ isomorphic to $C_p \times C_p$ (p prime) was obtained, using only the properties of the series $N(g, h)$. For related work, see Chapter 8 of [**57**].

5.3. Other finite groups. It is natural to ask about Moonshine for other groups. Indeed, the Hauptmodul for $\Gamma_0(2)+$ looks like

$$q^{-1} + 4372q + 96256q^2 + 12\,40002q^3 + \cdots \tag{5.5a}$$

and we find the relations

$$4372 = 4371 + 1 \ , \qquad 96256 = 96255 + 1 \ , \qquad 12\,40002 = 11\,39374 + 4371 + 2 \cdot 1 \ , \tag{5.5b}$$

where 1, 4371, 96255, and 11 39374 are all dimensions of irreducible representations of the Baby Monster $\mathbb{B}$. Thus we find Moonshine for $\mathbb{B}$! We will return to this example shortly.

Of course any subgroup of $\mathbb{M}$ automatically inherits Moonshine by restriction, but obviously this isn't interesting. Most constructions of the Leech lattice start with Mathieu's sporadic M_{24} (see e.g. Chapters 10 and 11 of [**26**]), and most constructions of the Monster involve the Leech lattice. Thus we are led to the following natural hierarchy of (most) sporadics:

(1) M_{24} (from which we can get M_{11}, M_{12}, M_{22}, M_{23}); which leads to
(2) $Co_0 = C_2 \times Co_1$ (from which we get HJ, HS, McL, Suz, Co_3, Co_2); which leads to
(3) $\mathbb{M}$ (from which we get He, Fi_{22}, Fi_{23}, Fi'_{24}, HN, Th, $\mathbb{B}$).

It can thus be argued that we could approach problems in Monstrous Moonshine, by first addressing in order M_{24} and Co_1, which should be much simpler. Indeed, the full VOA orbifold theory — i.e. the complete analogue of §5.1 — for M_{24} has been established in [**38**] (the relevant series $\mathcal{Z}(g,h)$ had already been constructed in [**88**]).

Largely by trial and error, Queen [**101**] established Moonshine for the following groups (all essentially centralisers of elements of $\mathbb{M}$): Co_0, Th, $3.2.Suz$, $2.HJ$, HN, $2.A_7$, He, M_{12} (by e.g. '$2.HJ$' we mean C_2 is normal and HJ is the quotient $2.HJ/C_2$). In particular, to each element g of these groups, there corresponds a series $Q_g(\tau) = q^{-1} + \sum_{n=0}^{\infty} a_n(g)q^n$, which is a Hauptmodul for some modular group of moonshine-type, and where each $g \mapsto a_n(g)$ is a virtual character. For Th, HN, He and M_{12} it is a proper character. Other differences with Monstrous Moonshine are that there can be a preferred nonzero value for the constant term a_0, and that although $\Gamma_0(N)$ will be a subgroup of the fixing group, it won't necessarily be normal. We will return to these results next section, where we will see that many seem to come out of the Moonshine for $\mathbb{M}$. About half of Queen's Hauptmoduls Q_g for Co_0 do not arise as a McKay–Thompson series for $\mathbb{M}$. Norton's conjectures in §5.1 are a reinterpretation and extension of Queen's work.

Queen never reached $\mathbb{B}$ because of its size. However, the Moonshine (5.5) for $\mathbb{B}$ falls into her and Norton's scheme because (5.5a) is the McKay–Thompson series associated to class 2A of $\mathbb{M}$, and the centraliser of an element in 2A is a double cover of $\mathbb{B}$.

There can't be a VOA $V = \oplus_n V_n$ with graded dimension (5.5a) and automorphisms in $\mathbb{B}$, because e.g. the $\mathbb{B}$-module V_3 doesn't contain V_2 as a submodule. However, Höhn deepened the analogy between $\mathbb{M}$ and $\mathbb{B}$ by constructing a vertex operator superalgebra $V\mathbb{B}^\natural$ of rank $c = 23.5$, called the *shorter Moonshine module*, closely related to $V^\natural$ (see e.g. [**56**]). Its automorphism group is $C_2 \times \mathbb{B}$. Just as $\mathbb{M}$ is the automorphism group of the Griess algebra $V_2^\natural$, so is $\mathbb{B}$ the automorphism group of the algebra $(V\mathbb{B}^\natural)_2$. Just as $V^\natural$ is associated to the Leech lattice Λ, so is $V\mathbb{B}^\natural$ associated to the shorter Leech lattice O_{23}, the unique 23-dimensional positive-definite self-dual lattice with no vectors of length 2 or 1 (see e.g. Chapter 6 of [**26**]). The automorphism group of O_{23} is $C_2 \times Co_2$.

There has been no interesting Moonshine rumoured for the remaining six sporadics (the *pariahs* J_1, J_3, Ru, ON, Ly, J_4). There will be some sort of Moonshine for any group which is an automorphism group of a vertex operator algebra (so this means any finite group [**37**]!). Many finite groups of Lie type should arise as automorphism groups of VOAs associated to affine algebras except defined over finite fields. But apparently all known examples of genus-0 Moonshine are limited to the groups involved with $\mathbb{M}$.

5.4. Modular Moonshine. Consider an element $g \in \mathbb{M}$. We expect from [**101**], [**97**], [**36**] that there is a Moonshine for the centraliser $C_{\mathbb{M}}(g)$ of g in $\mathbb{M}$, governed by the g-

twisted module $V^\natural(g)$. Unfortunately, $V^\natural(g)$ is not usually itself a VOA, so the analogy with $\mathbb{M}$ is not perfect. Ryba found it interesting that, for $g \in \mathbb{M}$ of prime order p, Norton's series $N(g,h)$ is a McKay–Thompson series (and has all the associated nice properties) whenever h is *p-regular* (i.e. h has order coprime to p). This special behaviour of p-regular elements suggested to him to look at modular representations.

The basics of modular representations and Brauer characters are discussed in sufficient detail in Chapter 2 of [**31**]. A *modular representation* ρ of a group G is a representation defined over a field of positive characteristic p dividing the order $|G|$ of G. Such representations possess many special (i.e. unpleasant) features. For one thing, they are no longer completely reducible (so the role of irreducible modules as direct summands will be replaced with their role as composition factors). For another, the usual notion of character (the trace of representation matrices) loses its usefulness and is replaced by the more subtle *Brauer character* $\beta(\rho)$: a complex-valued class function on $\mathbb{M}$ which is only well-defined on the p-regular elements of G.

THEOREM 7. [**105**], [**17**], [**13**] *Let $g \in \mathbb{M}$ be any element of prime order p, for any p dividing $|\mathbb{M}|$. Then there is a vertex operator superalgebra ${}^g\mathcal{V} = \oplus_{n\in\mathbb{Z}}{}^g\mathcal{V}_n$ defined over the finite field $\mathbb{F}_p$ and acted on by the centraliser $C_\mathbb{M}(g)$. If $h \in C_\mathbb{M}(g)$ is p-regular, then the graded Brauer character*

$$R(g,h;\tau) := q^{-1}\sum_{n\in\mathbb{Z}} \beta({}^g\mathcal{V}_n)(h)\, q^n$$

equals the McKay–Thompson series $T_{gh}(\tau)$. Moreover, for g belonging to any conjugacy class in $\mathbb{M}$ except 2B, 3B, 5B, 7B, or 13B, this is in fact an ordinary VOA (i.e. the 'odd' part vanishes), while in the remaining cases the graded Brauer characters of both the odd and even parts can separately be expressed using McKay–Thompson series.

By a vertex operator superalgebra, we mean there is a $\mathbb{Z}_2$-grading into even and odd subspaces, and for u, v both odd the commutator in (4.2a) is replaced by an anticommutator. In the proof, the superspaces arise as cohomology groups, which naturally form an alternating sum. The centralisers $C_\mathbb{M}(g)$ in the Theorem are quite nice: e.g. for g in classes 2A, 2B, 3A, 3B, 3C, 5A, 5B, 7A, 11A, respectively, these involve the sporadic groups $\mathbb{B}$, Co_1, Fi'_{24}, Suz, Th, HN, HJ, He, and M_{12}. The proof for $p = 2$ is not complete at the present time. The conjectures in [**105**] concerning modular analogues of the Griess algebra for several sporadics follow from Theorem 7.

Can these modular ${}^g\mathcal{V}$s be interpreted as a reduction mod p of (super)algebras in characteristic 0? Also, what about elements g of composite order?

CONJECTURE 8. [**13**] *Choose any $g \in \mathbb{M}$ and let n denote its order. Then there is a $\frac{1}{n}\mathbb{Z}$-graded superspace ${}^g\widehat{\mathcal{V}} = \oplus_{i\in\frac{1}{n}\mathbb{Z}}{}^g\widehat{\mathcal{V}}_i$ over the ring of cyclotomic integers $\mathbb{Z}[e^{2\pi i/n}]$. It is often (but probably not always) a vertex operator superalgebra — in particular ${}^1\widehat{\mathcal{V}}$ is an integral form of the Moonshine module $V^\natural$. Each ${}^g\widehat{\mathcal{V}}$ carries a representation of a central extension of $C_\mathbb{M}(g)$ by C_n. Define the graded trace*

$$B(g,h;\tau) = q^{-1}\sum_{i\in\frac{1}{n}\mathbb{Z}} \mathrm{ch}_{{}^g\widehat{\mathcal{V}}_i}(h)\, q^i\ .$$

If $g, h \in \mathbb{M}$ commute and have coprime orders, then $B(g, h; \tau) = T_{gh}(\tau)$. If all q-coefficients of T_g are nonnegative, then the 'odd' part of ${}^g\widehat{\mathcal{V}}$ vanishes, and ${}^g\widehat{\mathcal{V}}$ is the g-twisted module $V^\natural(g)$ of [**36**]. *If g has prime order p, then the reduction mod p of ${}^g\widehat{\mathcal{V}}$ is the modular vertex operator superalgebra ${}^g\mathcal{V}$ of Theorem 7.*

When we say ${}^1\widehat{\mathcal{V}}$ is an integral form for $V^\natural$, we mean that ${}^1\widehat{\mathcal{V}}$ has the same structure as a VOA, with everything defined over $\mathbb{Z}$, and tensoring it with $\mathbb{C}$ recovers $V^\natural$. This remarkable conjecture, which tries to explain Theorem 7, is completely open.

5.5. The geometry of Moonshine. Algebra is the mathematics of structure, and so of course it has a profound relationship with every area of mathematics. Therefore the trick for finding possible fingerprints of Moonshine in say geometry is to look there for modular functions. And that search quickly leads to the elliptic genus.

For details see e.g. [**55**], [**108**], [**112**]. All manifolds here are compact, oriented and differentiable. In Thom's cobordism ring Ω, elements are equivalence classes of cobordant manifolds, addition is connected sum, and multiplication is Cartesian product. The *universal elliptic genus* $\phi(M)$ is a ring homomorphism from $\mathbb{Q} \otimes \Omega$ to the ring of power series in q, which sends n-dimensional manifolds with spin connections to a weight $n/2$ modular form of $\Gamma_0(2)$ with integer coefficients. Several variations and generalisations have been introduced, e.g. the Witten genus assigns spin manifolds with vanishing first Pontryagin class a weight $n/2$ modular form of $\mathrm{SL}_2(\mathbb{Z})$ with integer coefficients.

Several deep relationships between elliptic genera and the general material reviewed elsewhere in this paper, have been uncovered. For instance, the important rigidity property of the Witten genus with respect to any compact Lie group action on the manifold, is a consequence of the modularity of the characters of affine algebras (our Theorem 4) [**81**]. The elliptic genus of a manifold M has been interpreted as the graded dimension of a vertex operator superalgebra constructed from M [**111**]. Seemingly related to this, [**18**] recovered the elliptic genus of a Calabi–Yau manifold X from the sheaf of vertex algebras in the chiral de Rham complex [**85**] attached to X. Unexpectedly, the elliptic genus of even-dimensional projective spaces P^{2n} has nonnegative coefficients and in fact equals the graded dimension of a certain vertex algebra [**86**]; this suggests interesting representation-theoretic questions in the spirit of Monstrous Moonshine. In physics, elliptic genera arise as partition functions of $N = 2$ superconformal field theories [**120**]. Mason's constructions [**88**] associated to Moonshine for the Mathieu group M_{24} have been interpreted as providing a geometric model ('elliptic system') for elliptic cohomology $\mathrm{Ell}^*(BM_{24})$ of the classifying space of M_{24} [**112**], [**39**]. The Witten genus (normalised by η^8) of the Milnor–Kervaire manifold M_0^8, an 8-dimensional manifold built from the E_8 diagram, equals $j^{\frac{1}{3}}$ [**55**] (recall (3.5)).

Hirzebruch's 'prize question' (p.86 of [**55**]) asks for the construction of a 24-dimensional manifold M with Witten genus J (after being normalised by η^{24}). We would like $\mathbb{M}$ to act on M by diffeomorphisms, and the twisted Witten genera to be the McKay–Thompson series T_g. It would also be nice to associate Norton's series $N(g, h)$ to this Moonshine manifold. Constructing such a manifold is perhaps the remaining Holy Grail of Monstrous Moonshine.

Hirzebruch's question was partially answered by Mahowald and Hopkins [**84**], who constructed a manifold with Witten genus J, but couldn't show it would support an effective action of $\mathbb{M}$. Related work is [**3**], who constructed several actions of $\mathbb{M}$ on e.g. 24-dimensional manifolds (but none of which could have genus J), and [**73**], who showed the graded dimensions of the subspaces $V^\natural_\pm$ of the Moonshine module are twisted $\widehat{A}$-genera of Milnor–Kervaire's manifold M_0^8 (the $\widehat{A}$-genus is the specialisation of elliptic genus to the cusp $\mathrm{i}\infty$).

There has been a second conjectured relationship between geometry and Monstrous Moonshine. *Mirror symmetry* says that most Calabi–Yau manifolds come in closely related pairs. Consider a 1-parameter family X_z of Calabi–Yau manifolds, with mirror X^* given by the resolution of an orbifold X/G for G finite and abelian. Then the Hodge numbers $h^{1,1}(X)$ and $h^{2,1}(X^*)$ will be equal, and more precisely the moduli space of (complexified) Kähler structures on X will be locally isometric to the moduli space of complex structures on X^*. The 'mirror map' $z(q)$, which can be defined using the Picard–Fuchs equation [**95**], gives a canonical map between those moduli spaces. For example, $x_1^4 + x_2^4 + x_3^4 + x_4^4 + z^{-1/4}x_1x_2x_3x_4 = 0$ is such a family of K3 surfaces, where $G = C_4 \times C_4$. Its mirror map is given by

$$z(q) = q - 104q^2 + 6444q^3 - 311744q^4 + 13018830q^5 - 493025760q^6 + \cdots \, . \tag{5.6}$$

Lian–Yau [**80**] noticed that the reciprocal $1/z(q)$ of the mirror map in (5.6) equals the McKay–Thompson series $T_g(\tau) + 104$ for g in class 2A of $\mathbb{M}$. After looking at several other examples with similar conclusions, they proposed their *Mirror-Moonshine Conjecture*: The reciprocal $1/z$ of the mirror map of a 1-parameter family of K3 surfaces with an orbifold mirror, will be a McKay–Thompson series (up to an additive constant).

A counterexample (and more examples) are given in §7 of [**118**]. In particular, although there are relations between mirror symmetry and modular functions (see e.g. [**51**] and [**54**]), there doesn't seem to be any special relation with the Monster. Doran [**40**] 'demystifies the Mirror-Moonshine phenomenon' by finding necessary and sufficient conditions for $1/z$ to be a modular function for a modular group commensurable with $\mathrm{SL}_2(\mathbb{Z})$.

6. The physics of Moonshine

The physical side (perturbative string theory, or equivalently conformal field theory) of Moonshine was noticed early on, and has profoundly influenced the development of Moonshine and VOAs. This is a very rich subject, which we can only superficially touch on. The book [**32**], with its extensive bibliography, provides an introduction but will be difficult reading for many mathematicians (as will this section!) The treatment in [**45**] is more accessible and shows how naturally VOAs arise from the physics. This effectiveness of physical interpretations isn't magic — it merely tells us that many of our finite-dimensional objects are seen much more clearly when studied through infinite-dimensional structures (often by being 'looped'). Of course Moonshine, which teaches us to study the finite group $\mathbb{M}$ via its infinite-dimensional module $V^\natural$, fits perfectly into this picture.

A conformal field theory (CFT) is a quantum field theory on 2-dimensional space-time, whose symmetries include the conformal transformations. In string theory the basic objects are finite curves ('strings') rather than points ('particles'), and the CFT lives on the surface traced by the strings as they evolve (colliding and separating) through time. Each CFT is associated with a pair $\mathcal{V}_L, \mathcal{V}_R$ of mutually commuting VOAs, called its *chiral algebras* [**6**]. For example, strings living on a compact Lie group manifold (the so-called Wess–Zumino–Witten model) will have chiral algebras given by affine algebra VOAs. The space $\mathcal{H}$ of states for the CFT carries a representation of $\mathcal{V}_L \otimes \overline{\mathcal{V}}_R$, and many authors have (somewhat optimisticly) concluded that the study of CFTs reduces to that of VOA representation theory. Rational VOAs correspond to the important class of *rational* CFTs, where $\mathcal{H}$ decomposes into a finite sum $\oplus M_L \otimes M_R$ of irreducible modules. The Virasoro algebra (4.3) arises naturally in CFT through infinitesimal conformal transformations. The vertex operator $Y(\phi, z)$, for the space-time parameter $z = e^{t+\mathrm{i}x}$, is the quantum field which creates from the vacuum $|0\rangle \in \mathcal{H}$ the state $|\phi\rangle \in \mathcal{H}$ at time $t = -\infty$: $|\phi\rangle = \lim_{z\to 0} Y(\phi, z)\,|0\rangle$. In particular, Borcherds' definition [**9**] of VOAs can be interpreted as an axiomatisation of the notion of chiral algebra in CFT, and for this reason alone is important.

In CFT, the Hauptmodul property of Moonshine is hard to interpret, and a less direct formulation like that in [**116**] is needed. However, both the statement and proof of Theorem 5 are natural from the CFT framework (see [**45**]) — e.g. the modularity of the series T_g and $N(g, h)$ are automatic in CFT. This modularity arises in CFT through the equivalence of the Hamiltonian formulation, which describes concretely the graded spaces we take traces on (and hence the coefficients of our q-expansions), and the Feynman path formalism, which interprets these graded traces as sections over moduli spaces (and hence makes modularity manifest). Beautiful reviews are sketched in [**119**], [**120**].

Because $V^\natural$ is so mathematically special, it may be expected that it corresponds to interesting physics. Certainly it has been the subject of some speculation. There will be a $c = 24$ rational CFT whose chiral algebra $\mathcal{V}_L$ and state space $\mathcal{H}$ are both $V^\natural$, while $\mathcal{V}_R$ is trivial (this is possible because $V^\natural$ is holomorphic). This CFT is nicely described in [**34**]; see also [**35**]. The Monster is the symmetry of that CFT, but the Bimonster $\mathbb{M} \wr C_2$ will be the symmetry of a rational CFT with $\mathcal{H} = V^\natural \otimes \overline{V^\natural}$. The paper [**27**] finds a family of D-branes for the latter theory which are in one-to-one correspondence with the elements of $\mathbb{M}$, and their 'overlaps' $\langle\!\langle g \| q^{\frac{1}{2}(L_0+\bar{L}_0-\frac{c}{24})} \| h\rangle\!\rangle$ equal the McKay–Thompson series $T_{g^{-1}h}$. However, we still lack any explanation as to why a CFT involving $V^\natural$ should yield interesting *physics*.

Almost every facet of Moonshine finds a natural formulation in CFT, where it often was discovered first. For example, the 'No-Ghost' Theorem of Brower–Goddard–Thorn was used to great effect in [**11**] to understand the structure of the Monster Lie algebra $\mathfrak{m}$. On a finite-dimensional manifold M, the index of the Dirac operator D in the heat kernel interpretation is a path integral in supersymmetric quantum mechanics, i.e an integral over the free loop space $\mathcal{L}M = \{\gamma : S^1 \to M\}$; the string theory version of this is that the index of the Dirac operator on $\mathcal{L}M$ should be an integral over $\mathcal{L}(\mathcal{L}M)$, i.e. over smooth maps of tori into M, and this is just the elliptic genus, and explains why it should be modular. The orbifold construction of [**36**] comes straight from CFT (although [**43**]'s construction of $V^\natural$ predates CFT orbifolds by a year and in fact influenced their development in physics). That said, the translation process from physics to mathematics of course is never easy — Borcherds' definition [**9**] is a prime example!

But from this standpoint, what is most exciting is what hasn't yet been fully exploited.

String theory tells us that CFT can live on any surface Σ. The VOAs, including the geometric VOAs of [**59**], capture CFT in genus 0. The graded dimensions and traces considered above concern CFT quantities ('conformal blocks') at genus 1: $\tau \mapsto e^{2\pi \mathrm{i}\tau}$ maps $\mathbb{H}$ onto a cylinder, and the trace identifies the two ends. But there are analogues of all this at higher genus [**123**] (though the formulas can rapidly become awkward). For example, the graded dimension of e.g. the $V^\natural$ CFT in genus 2 is computed in [**117**], and involves e.g. Siegel theta functions. The orbifold theory in §5.1 is genus 1: each 'sector' (g, h) corresponds to a homomorphism from the fundamental group $\mathbb{Z}^2$ of the torus into the orbifold group G (e.g $G = \mathbb{M}$) — g and h are the targets of the two generators of $\mathbb{Z}^2$ and hence must commute. More generally, the sectors will correspond to each homomorphism $\varphi : \pi_1(\Sigma) \to G$, and to each we will get a higher genus trace $\mathcal{Z}(\varphi)$, which will be a function on the Teichmüller space T_g (generalising the upper half-plane $\mathbb{H}$ for genus 1). The action of $\mathrm{SL}_2(\mathbb{Z})$ on the $N(g, h)$ generalises to the action of the mapping class group on $\pi_1(\Sigma)$ and T_g. See e.g. [**4**] for some thoughts in this direction.

7. Conclusion

There are different basic aspects to Monstrous Moonshine: (i) why modularity enters at all; (ii) why in particular we have genus 0; and (iii) what does it have to do with the Monster. We understand (i) best. There will be a Moonshine-like relation between any (subgroup of the) automorphism group of any rational VOA, and the characters χ_M, and the same can be expected to hold of the orbifold characters $\mathcal{Z}$ in §5.1.

To prove the genus 0 property of the T_g, we needed recursions obtained one way or another from the Monster Lie algebra $\mathfrak{m}$, and from these we apply Theorem 6. These recursions are very special, but so presumably is the genus 0 property. The suggestion of [**20**] though is that we may be able to considerably simplify this part of the argument.

Every group known to have rich Moonshine properties is contained in the Monster. Our understanding of this seemingly central role of $\mathbb{M}$ is the poorest of those three aspects.

It should be clear from this review, of the central role VOAs play in our current understanding of Moonshine. The excellent review [**39**] makes this point even more forcefully. It can be (and has been) questioned though whether the full and difficult machinery of VOAs is really needed to understand this, i.e. whether we really have isolated the key conjunction of properties needed for Moonshine to arise. CFT has been an invaluable guide thus far, but perhaps we are a little too steeped in its lore.

Moonshine (in its more general sense) is a relation between algebra and number theory, and its impact on algebra has been dramatic (e.g. VOAs, $V^\natural$, Borcherds–Kac–Moody algebras). Its impact on number theory has been far less so. This may merely be a temporary accident due to the backgrounds of most researchers (including the mathematical physicists) working to date in the area. But the most exciting prospects for the future of Moonshine (in this writer's opinion) are in the direction of number theory. Hints of this future can be found in e.g. [**121**], [**41**], [**14**], [**33**], [**52**], [**94**].

Acknowledgements. This paper was written at IHES and at U of Wales Swansea, and I thank both for their hospitality. Much of what I've learned of Moonshine came from conversations over the years with Chris Cummins and John McKay, and I warmly thank both. I appreciate the many comments and suggestions provided by the readers of the first draft; in this vein I would particularly like to mention James Lepowsky and Fyodor Malikov. My research is supported in part by NSERC.

References[†]

1. D. ALEXANDER, C. CUMMINS, J. MCKAY, and C. SIMONS, 'Completely replicable functions', *Groups, Combinatorics and Geometry* (ed. M.W. Liebeck and J. Saxl, Cambridge Univ. Press, 1992) 87–98.

2. O. ALVAREZ, 'Conformal anomalies and the index theorem', *Nucl. Phys.* B286 (1987) 175–188.

3. M.G. ASCHBACHER, 'Finite groups acting on homology manifolds', *Olga Taussky-Todd: in Memoriam, Pacific J. Math. Special Issue* (ed. M. Aschbacher et al, Pacific Journal of Mathematics, Berkeley, 1997) 3–36.

4. P. BÁNTAY, 'Higher genus moonshine', *Moonshine, the Monster, and Related Topics, Contemp. Math.* 193 (Amer. Math. Soc., Providence, 1996) 1–8.

5. P. BÁNTAY, 'The kernel of the modular representation and the Galois action in RCFT', *Commun. Math. Phys.* 233 (2003) 423–438.

6. A. BELAVIN, A.M. POLYAKOV, and A.B. ZAMOLODCHIKOV, 'Infinite conformal symmetries in two-dimensional quantum field theory', *Nucl. Phys.* **B241** (1984) 333–380.

7. J.S. BIRMAN, *Braids, Links and Mapping Class Groups* (Princeton Univ. Press, 1974).

8. S. BLOCH, 'Zeta values and differential operators on the circle', *J. Alg.* 182 (1996) 476–500.

9. R.E. BORCHERDS, 'Vertex algebras, Kac–Moody algebras, and the Monster', *Proc. Natl. Acad. Sci. (USA)* 83 (1986) 3068–3071.

10. R.E. BORCHERDS, 'Generalized Kac–Moody Lie algebras', *J. Alg.* 115 (1988) 501–512.

11. R.E BORCHERDS, 'Monstrous moonshine and monstrous Lie superalgebras', *Invent. Math.* 109 (1992) 405–444.

12. R.E BORCHERDS, 'Sporadic groups and string theory', *First European Congress of Math. Paris (1992)*, Vol.I (Birkhäuser, Basel, 1994) 411–421.

13. R.E. BORCHERDS, 'Modular moonshine III', *Duke Math. J.* 93 (1998) 129–154.

14. R.E. BORCHERDS, 'Automorphic forms with singularities on Grassmannians', *Invent. Math.* 132 (1998) 491–562.

15. R.E. BORCHERDS, 'What is Moonshine?', *Proc. Intern. Congr. Math. 1998 (Berlin)*, Vol. 1 (Documenta Mathematica, Bielefeld, 1998) 607–615.

16. R.E. BORCHERDS, 'Problems in Moonshine', *First Intern. Congr. Chinese Math.* (Amer. Math. Soc., Providence, 2001) 3–10.

17. R.E. BORCHERDS and A.J.E. RYBA, 'Modular moonshine II', *Duke Math. J.* 83 (1996) 435–459.

18. L.A. BORISOV and A. LIBGOBER, 'Elliptic genera of toric varieties and applications to mirror symmetry', *Invent. Math.* 140 (2000) 453–485.

[†] For a more comprehensive Moonshine bibliography, especially for older papers, see http://cicma.mathstat.concordia.ca/faculty/cummins/moonshine.html

19. S.-P. Chan, M.-L. Lang, and C.-H. Lim, 'Some modular functions associated to Lie algebra E_8', *Math. Z.* 211 (1992) 223–246.

20. H. Cohn and J. McKay, 'Spontaneous generation of modular invariants', *Math. of Comput.* 65 (1996) 1295–1309.

21. J.H. Conway, 'Monsters and Moonshine', *Math. Intell.* 2 (1980) 165–171.

22. J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker and R.A. Wilson, *An Atlas of Finite Groups* (Clarendon Press, Oxford, 1985).

23. J.H. Conway, J. McKay and A. Sebbar, 'On the discrete groups of moonshine', *Proc. Amer. Math. Soc.* (to appear).

24. J.H. Conway and S.P. Norton, 'Monstrous Moonshine', *Bull. London Math. Soc.* 11 (1979) 308–339.

25. J.H. Conway, S.P. Norton, and L.H. Soicher, 'The Bimonster, the group Y_{555}, and the projective plane of order 3', *Computers in Algebra* (Dekker, New York, 1988) 27–50.

26. J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd edn (Springer, Berlin, 1999).

27. B. Craps, M.R. Gaberdiel, J.A. Harvey, 'Monstrous branes', *Commun. Math. Phys.* 234 (2003) 229–251.

28. C.J. Cummins, 'Congruence subgroups of groups commensurable with $\mathrm{PSL}(2,\mathbb{Z})$ of genus 0 and 1', Preprint.

29. C.J. Cummins and T. Gannon, 'Modular equations and the genus zero property of moonshine functions', *Invent. Math.* 129 (1997) 413–443.

30. C.J. Cummins and S.P. Norton, 'Rational Hauptmodul are replicable', *Canad. J. Math.* 47 (1995) 1201–1218.

31. C.W. Curtis and I. Reiner, *Methods of Representation Theory with Applications to Finite Groups and Orders*, Vol.I (Wiley, New York, 1981).

32. P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory* (Springer, New York, 1997).

33. R. Dijkgraaf, 'The mathematics of fivebranes', *Proc. Intern. Congr. Math. 1998 (Berlin)*, Vol.III (Documenta Mathematica, Bielefeld, 1998) 133–142.

34. L. Dixon, P. Ginsparg and J.A. Harvey, 'Beauty and the beast: superconformal symmetry in a monster module', *Commun. Math. Phys.* 119 (1988) 221–241.

35. L. Dolan, P. Goddard and P. Montague, 'Conformal field theory of twisted vertex operators', *Nucl. Phys.* B338 (1990) 529–601.

36. C. Dong, H. Li and G. Mason, 'Modular invariance of trace functions in orbifold theory and generalized moonshine', *Commun. Math. Phys.* 214 (2000) 1–56.

37. C. Dong and G. Mason, 'Nonabelian orbifolds and the boson-fermion correspondence', *Commun. Math. Phys.* 163 (1994) 523–559.

38. C. Dong and G. Mason, 'An orbifold theory of genus zero associated to the sporadic group M_{24}', *Commun. Math. Phys.* 164 (1994) 87–104.

39. C. Dong and G. Mason, 'Vertex operator algebras and moonshine: A survey', *Progress in Algebraic Combinatorics, Adv. Stud. Pure Math.* 24 (Math. Soc. Japan, Tokyo, 1996) 101–136.

40. C.F. Doran, 'Picard–Fuchs uniformization and modularity of the mirror map', *Commun. Math. Phys.* 212 (2000) 625–647.

41. V.G. Drinfeld, 'On quasitriangular quasi-Hopf algebras on a group that is closely related with Gal $\overline{\mathbb{Q}}/\mathbb{Q}$', *Leningrad. Math. J.* 2 (1991) 829–860.

42. I. Frenkel, Y.-Z. Huang and J. Lepowsky, 'On Axiomatic Approaches to Vertex Operator Algebras and Modules', *Mem. Amer. Math. Soc.* 104 (1993) 1–63.

43. I. Frenkel, J. Lepowsky and A. Meurman, 'A natural representation of the Fischer–Griess monster with the modular function J as character', *Proc. Natl. Acad. Sci. USA* 81 (1984) 3256–3260.

44. I. Frenkel, J. Lepowsky and A. Meurman, *Vertex Operator Algebras and the Monster* (Academic Press, New York, 1988).

45. M.R. Gaberdiel and P. Goddard, 'Axiomatic conformal field theory', *Commun. Math. Phys.* 209 (2000) 549–594.

46. T. Gannon, *Moonshine Beyond the Monster* (Cambridge Univ. Press, to appear).

47. G. Glauberman and S.P. Norton, 'On McKay's connection between the affine E_8 diagram and the Monster', *Proc. on Moonshine and Related Topics* (Amer. Math. Soc., Providence, 2001) 37–42.

48. P. Goddard, 'The work of Richard Ewen Borcherds', *Proc. Intern. Congr. Math. 1998 (Berlin)*, Vol. I (Documenta Mathematica, Bielefeld, 1998) 99–108.

49. D. Gorenstein, *Finite Simple Groups: An Introduction to their Classification* (Plenum, New York, 1982).

50. R. Griess, 'The friendly giant', *Invent. Math.* 68 (1982) 1–102.

51. V.A. Gritsenko and V.V. Nikulin, 'The arithmetic mirror symmetry and Calabi–Yau manifolds', *Commun. Math. Phys.* 210 (2000) 1–11.

52. S. Gukov and C. Vafa, 'Rational conformal field theories and complex multiplication', Preprint (arXiv: hep-th/0203213).

53. K. Harada, M. Miyamoto, and H. Yamada, 'A generalization of Kac–Moody algebras', *Groups, Difference Sets, and the Monster* (de Gruyter, Berlin, 1996) 377–408.

54. J.A. Harvey and G. Moore, 'Algebras, BPS states, and strings', *Nucl. Phys.* B463 (1996) 315–368.

55. F. Hirzebruch, T. Berger, and R. Jung, *Manifolds and Modular Forms* (Friedr. Vieweg & Sohn, 1991).

56. G. Höhn, 'The group of symmetries of the shorter Moonshine module', Preprint (arXiv:math.QA/0210076).

57. T. Hsu, *Quilts: Central Extensions, Braid Actions, and Finite Groups*, *Lecture Notes in Mathematics 1731* (Springer, Berlin, 2000).

58. Y.-Z. Huang, 'Binary trees and finite-dimensional Lie algebras', *Algebraic Groups and their Generalizations: Quantum and Infinite-Dimensional Methods*, Proc. Symp. Pure Math. 56, pt. 2 (Amer. Math. Soc., Providence, 1994) 337–348.

59. Y.-Z. Huang, *Two-dimensional Conformal Geometry and Vertex Operator Algebras* (Birkhäuser, Boston, 1997).

60. A.A. Ivanov, 'Geometric presentations of groups with an application to the Monster', *Proc. Intern. Congr. Math. 1990 (Kyoto)*, Vol. II (Springer, Hong Kong, 1991) 1443–

1453.

61. A.A. Ivanov, 'Y-groups via transitive extension', *J. Alg.* 218 (1999) 412–435.

62. R. Ivanov and M.P. Tuite, 'Some irrational generalized Moonshine from orbifolds', *Nucl. Phys.* B635 (2002) 473–491.

63. E. Jurisich, 'An exposition of generalized Kac–Moody algebras', *Contemp. Math.* 194 (Amer. Math. Soc., Providence, 1996) 121–159.

64. E. Jurisich, J. Lepowsky and R.L. Wilson, 'Realizations of the Monster Lie algebra', *Selecta Math. (NS)* 1 (1995) 129–161.

65. V.G. Kac, 'An elucidation of: Infinite-dimensional algebras, Dedekind's η-function, classical Möbius function and the very strange formula. $E_8^{(1)}$ and the cube root of the modular invariant j', *Adv. in Math.* 35 (1980) 264–273.

66. V.G. Kac, *Infinite Dimensional Lie Algebras*, 3rd edn, (Cambridge Univ. Press, 1990).

67. V.G. Kac, *Vertex Operators for Beginners*, Univ. Lecture Series, Vol. 10 (Amer. Math. Soc., Providence, 1997).

68. V.G. Kac and S.-J. Kang, 'Trace formula for graded Lie algebras and Monstrous Moonshine', *Representations of Groups, Banff (1994)* (Amer. Math. Soc., Providence, 1995) 141–154.

69. V.G. Kac and D.H. Peterson, 'Infinite dimensional Lie algebras, theta functions, and modular forms', *Adv. in Math.* 53 (1984) 125–264.

70. S. Kass, R.V. Moody, J. Patera and R. Slansky, *Affine Lie Algebras, Weight Multiplicities, and Branching Rules*, Vol. 1 (Univ. California Press, Berkeley, 1990).

71. M. Knopp and G. Mason, 'Generalized modular forms', *J. Number Theory* 99 (2003) 1–28.

72. D.N. Kozlov, 'On completely replicable functions and extremal poset theory', MSc thesis, Univ. of Lund, Sweden, 1994.

73. R. Kultze, 'Elliptic genera and the moonshine module', *Math. Z.* 223 (1996) 463–471.

74. M.-L. Lang, 'On a question raised by Conway–Norton', *J. Math. Soc. Japan* 41 (1989) 263–284.

75. S. Lang, *Elliptic Functions*, 2nd edn (Springer, New York, 1997).

76. J. Lepowsky, 'Euclidean Lie algebras and the modular function j', *The Santa Cruz Conference on Finite Groups, Proc. Sympos. Pure Math.* 37 (Amer. Math. Soc., Providence, 1980) 567–570.

77. J. Lepowsky, 'Vertex operator algebras and the zeta function', *Recent Developments in Quantum Affine Algebras and Related Topics, Contemp. Math.* 248 (Amer. Math. Soc., Providence, 1999) 327–340.

78. J. Lepowsky, 'The work of Richard E. Borcherds', *Notices Amer. Math. Soc.* 46 (1999) 17–19.

79. J. Lepowsky and H. Li, *Introduction to Vertex Operator Algebras and their Representations* (Birkhäuser, Boston, 2004).

80. B.H. Lian and S.-T. Yau, 'Arithmetic properties of mirror map and quantum coupling', *Commun. Math. Phys.* 176 (1996) 163–191.

81. B.H. Lian and G.J. Zuckerman, 'Moonshine cohomology', *Moonshine and Vertex Operator Algebras* (Surikaisekikenkyusho Kokyuroku No. 904, Kyoto, 1995) 87–115.

82. K. Liu, 'On modular invariance and rigidity theorems', *J. Diff. Geom.* 41 (1995) 343–396.

83. I.G. Macdonald, 'Affine root systems and Dedekind's η-function', *Invent. Math.* 15 (1972) 91–143.

84. M. Mahowald and M. Hopkins, 'The structure of 24 dimensional manifolds having normal bundles which lift to BO[8]', *Recent Progress in Homotopy Theory, Contemp. Math.* 293 (Amer. Math. Soc., Providence, 2002) 89–110.

85. F. Malikov, V. Schechtman and A. Vaintrob, 'Chiral de Rham complex', *Commun. Math. Phys.* **204** (1999) 439–473.

86. F. Malikov and V. Schechtman, 'Deformations of vertex algebras, quantum cohomology of toric varieties, and elliptic genus', *Commun. Math. Phys.* 234 (2003) 77–100.

87. G. Mason, 'Finite groups and modular functions', *The Arcata Conference on Representations of Finite Groups, Proc. Sympos. Pure Math.* 47 (Amer. Math. Soc., Providence, 1987) 181–209.

88. G. Mason, 'On a system of elliptic modular forms attached to the large Mathieu group', *Nagoya Math. J.* 118 (1990) 177–193.

89. J. McKay, 'Graphs, singularities, and finite groups', *The Santa Cruz Conference on Finite Groups, Proc. Sympos. Pure Math.* 37 (Amer. Math. Soc., Providence, 1980) 183–186.

90. J. McKay, 'The essentials of Monstrous Moonshine', *Groups and Combinatorics – in Memory of M. Suzuki, Adv. Studies in Pure Math.* 32 (Math. Soc. Japan, Tokyo, 2001) 347–353.

91. J. McKay and H. Strauss, 'The q-series of monstrous moonshine and the decomposition of the head characters', *Commun. Alg.* 18 (1990) 253–278.

92. H. McKean and V. Moll, *Elliptic Curves: Function Theory, Geometry, Arithmetic* (Cambridge Univ. Press, 1999).

93. M. Miyamoto, '21 involutions acting on the Moonshine module', *J. Alg.* 175 (1995) 941–965.

94. G.W. Moore, 'Les Houches lectures on strings and arithmetic', Preprint (arXiv: hep-th/0401049).

95. D.R. Morrison, 'Picard–Fuchs equations and mirror maps for hypersurfaces', *Essays on Mirror Manifolds* (ed. S.-T. Yau, Intern. Press, Hong Kong, 1992) 241–264.

96. S.P. Norton, 'More on Moonshine', In: *Computational Group Theory* (ed. M.D. Atkinson, Academic Press, New York, 1984) 185–193.

97. S.P. Norton, 'Generalized moonshine', *The Arcata Conference on Representations of Finite Groups, Proc. Sympos. Pure Math.* 47 (Amer. Math. Soc., Providence, 1987) 208–209.

98. S.P. Norton, 'Constructing the Monster', *Groups, Combinatorics, and Geometry* (ed. M.W. Liebeck and J. Saxl, Cambridge Univ. Press, 1992) 63–76.

99. S.P. Norton, 'From moonshine to the Monster', *Proc. on Moonshine and Related Topics* (Amer. Math. Soc, Providence, 2001) 163–171.

100. A. Ogg, *Modular forms and Dirichlet series* (W. A. Benjamin, New York, 1969).

101. L. Queen, 'Modular functions arising from some finite groups', *Math. of Comput.* 37 (1981) 547–580.

102. U. Ray, 'Generalized Kac–Moody algebras and some related topics', *Bull. Amer. Math. Soc.* 38 (2000) 1–42.

103. W.F. Reynolds, 'Thompson's characterization of characters and sets of primes', *J. Alg.* 156 (1993) 237–243.

104. S.-S. Roan, 'Mirror symmetry of elliptic curves and Ising model', *J. Geom. Phys.* 20 (1996) 273–296.

105. A.J.E. Ryba, 'Modular moonshine?', *Moonshine, the Monster, and Related Topics, Contemp. Math.* 193 (Amer. Math. Soc., Providence, 1996) 307–336.

106. A.N. Schellekens, 'Meromorphic $c = 24$ conformal field theories', *Commun. Math. Phys.* 153 (1993) 159–185.

107. G. Segal, 'The definition of conformal field theory', *Differential Geometric Methods in Theoretical Physics* (ed. K. Bleuler and M. Werner, Academic Press, Boston, 1988) 165–171.

108. G. Segal, 'Elliptic cohomology', *Séminaire Bourbaki 1987-88, no. 695, Astérisque* 161-162 (1988) 187–201.

109. G.W. Smith, 'Replicant powers for higher genera', *Moonshine, the Monster, and Related Topics, Contemp. Math.* 193 (Amer. Math. Soc., Providence, 1996) 337–352.

110. S.D. Smith, 'On the head characters of the monster simple group', *Finite Groups – Coming of Age* (ed. J. McKay, Amer. Math. Soc., Providence, 1985) 303–313.

111. H. Tamanoi, *Elliptic Genera and Vertex Operator Superalgebras, Lecture Notes in Mathematics 1704* (Springer, Berlin, 1999).

112. C.B. Thomas, *Elliptic Cohomology* (Kluwer, New York, 1999).

113. J.G. Thompson, 'Finite groups and modular functions', *Bull. London Math. Soc.* 11 (1979) 347–351.

114. J.G. Thompson, 'Some numerology between the Fischer–Griess Monster and the elliptic modular function', *Bull. Lond. Math. Soc.* 11 (1979) 352–353.

115. J.G. Thompson, 'A finiteness theorem for subgroups of PSL(2,$\mathbb{R}$) which are commensurable with PSL$(2, \mathbb{Z})$', *Santa Cruz Conference on Finite Groups, Proc. Symp. Pure Math.* 37 (Amer. Math. Soc., Providence, 1980) 533–555.

116. M.P. Tuite, 'On the relationship between Monstrous Moonshine and the uniqueness of the Moonshine module', *Commun. Math. Phys.* 166 (1995) 495–532.

117. M.P. Tuite, 'Genus two meromorphic conformal field theories', *Proc. on Moonshine and Related Topics* (Amer. Math. Soc., Providence, 2001) 231–251.

118. H. Verrill and N. Yui, 'Thompson series and the mirror maps of pencils of K3 surfaces', *The Arithmetic and Geometry of Algebraic Cycles* (ed. B. Gordon et al, Centre Res. Math. Proc. and Lecture notes 24, 2000) 399–432.

119. E. Witten, 'Physics and geometry', *Proc. Intern. Congr. Math. 1986 (Berkeley)* (Amer. Math. Soc., Providence, 1987).

120. E. Witten, 'Elliptic genera and quantum field theory', *Commun. Math. Phys.* 109 (1987) 525–536.

121. E. Witten, 'Quantum field theory, Grassmannians, and algebraic curves', *Commun. Math. Phys.* 113 (1988) 529–600.

122. E. Witten, ‘Geometry and quantum field theory’, *Mathematics Into the Twenty-first Century*, Vol.II (Amer. Math. Soc., Providence, 1992) 479–491.

123. Y. Zhu, ‘Global vertex operators on Riemann surfaces’, *Commun. Math. Phys.* 165 (1994) 485–531.

124. Y. Zhu, ‘Modular invariance of characters of vertex operator algebras’, *J. Amer. Math. Soc.* 9 (1996) 237–302.

Terry Gannon
Department of Mathematics
University of Alberta
Edmonton, CANADA T6G 2G1

tgannon@math.ualberta.ca

arXiv:1211.5531v2 [math.RT] 15 Mar 2013

Much ado about Mathieu

TERRY GANNON
Department of Mathematics, University of Alberta,
Edmonton, Alberta, Canada T6G 2G1
e-mail: `tgannon@math.ualberta.ca`

March 18, 2013

Abstract

Eguchi, Ooguri and Tachikawa have observed that the elliptic genus of type II string theory on K3 surfaces appears to possess a Moonshine for the largest Mathieu group. Subsequent work by several people established a candidate for the elliptic genus twisted by each element of M_{24}. In this paper we prove that the resulting sequence of class functions are true characters of M_{24}, proving the Eguchi-Ooguri-Tachikawa conjecture. The integrality of multiplicities is proved using a small generalisation of Sturm's Theorem, while positivity involves a modification of a method of Hooley. We also prove the evenness property of the multiplicities, as conjectured by several authors. We also identify the role group cohomology plays in both K3-Mathieu Moonshine and Monstrous Moonshine; in particular this gives a cohomological interpretation for the non-Fricke elements in Norton's Generalised Monstrous Moonshine conjecture. We investigate the proposal of Gaberdiel-Hohenegger-Volpato that K3-Mathieu Moonshine lifts to the Conway group Co_1.

Contents

1 Introduction 2

2 K3-Mathieu Moonshine: Review 5

3 Weak K3-Mathieu Moonshine I: Integrality 11

4 Weak Mathieu Moonshine II: Positivity 17

5 Is the Conway group the stringy symmetry? 30

6 Speculations 33

References 37

1 Introduction

The elliptic genus (a.k.a. partition function) of a nonlinear sigma model with K3 target space is a very special function. On general grounds, it is a weak Jacobi form of index one and weight zero for $\mathbb{Z}^2 \rtimes \mathrm{SL}_2(\mathbb{Z})$, and is therefore equal to the unique (up to scaling) such function, $\phi_{0,1}(z,\tau)$. This implies it is topological in the sense that it is independent of where you are on the 20-(complex) dimensional K3 moduli space, or indeed of where you are on the 40-dimensional moduli space of K3 sigma models.

The world-sheet description of these string theories is as an $N = 4$ superconformal field theory, and thus the whole state space can be organised as an $N = 4$ superconformal representation. In particular, the elliptic genus can be decomposed into sums of elliptic genera of $N = 4$ superconformal representations. Eguchi, Ooguri and Tachikawa [24] observed that the multiplicities with which these $N = 4$ elliptic genera contribute to the elliptic genus are dimensions of M_{24}-group representations.

According to Conway, the sporadic simple group M_{24} is the most remarkable of all finite groups (p.300, [9]), with a great wealth of structure and applications. For instance it is the automorphism group of the Golay code, a stabilizer in the Leech lattice, and all symplectic symmetries of K3 surfaces can be embedded in it (in fact in a subgroup M_{23}) [42, 39]. Its order (size) is 244 823 040, and class number (i.e. number of conjugacy classes) is 26. The character table is given in Table 1; we retain the order of irreps(= irreducible representations) in [8], relate with a bar the complex conjugate irrep (when nonisomorphic), and write $\rho_0 = 1$ for the trivial representation. We write there $\alpha_t^s = (s + t\mathrm{i}\sqrt{7})/2$, $\beta_t^s = (s + t\mathrm{i}\sqrt{15})/2$, and $\gamma_t^s = (s + t\mathrm{i}\sqrt{23})/2$, for all choices of signs $s, t \in \{\pm 1\}$.

K3-Mathieu Moonshine was pushed further — indeed, made well-defined — by the work of Cheng [5], Gaberdiel, Hohenegger & Volpato [28, 29] and Eguchi & Hikami [20] who calculated the analogue $\phi_g(\tau, z)$ of the McKay-Thompson series here. The resulting functions are weak Jacobi forms for $\mathbb{Z}^2 \rtimes \Gamma_0(|g|)$ up to certain phases ($|g|$ denotes the order of $g \in M_{24}$). At the conclusion of this paper (see also [31]) we identify the cohomological source of these phases. The K3-Mathieu Moonshine fits beautifully into this general cohomological framework of conformal field theory (CFT) [31].

Once we know the weak Jacobi forms ϕ_g, we obtain modular forms f_g from (2.10) below and can read off the class functions H_n from (2.11) and (2.13) (a class function is a function constant on conjugacy classes). These H_n can also be obtained explicitly from (4.4) below. In any case, our most important result is the proof that these class functions H_n are all true representations of M_{24}. This can be regarded as a proof of the weak form of the K3-Mathieu moonshine conjecture.

Theorem A. *Each H_n ($n \geq 1$) is a true character of M_{24} (i.e. a sum of irreducible characters of M_{24}).*

Table 1. Character table of M_{24}

K_g	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	23A	23B
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ρ_1	23	7	-1	5	-1	-1	3	-1	3	1	-1	2	2	1	-1	1	-1	-1	0	0	0	0	-1	-1	0	0
ρ_2	45	-3	5	0	3	-3	1	1	0	0	-1	α_+^-	α_-^-	-1	0	1	0	1	α_-^+	α_+^+	0	0	α_+^-	α_-^-	-1	-1
$\overline{\rho_2}$	45	-3	5	0	3	-3	1	1	0	0	-1	α_-^-	α_+^-	-1	0	1	0	1	α_+^+	α_-^+	0	0	α_-^-	α_+^-	-1	-1
ρ_3	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	β_+^-	β_-^-	0	0	1	1
$\overline{\rho_3}$	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	β_-^-	β_+^-	0	0	1	1
ρ_4	252	28	12	9	0	4	4	0	2	1	0	0	0	0	2	-1	1	0	0	0	-1	-1	0	0	-1	-1
ρ_5	253	13	-11	10	1	-3	1	1	3	-2	1	1	1	-1	-1	0	0	1	-1	-1	0	0	1	1	0	0
ρ_6	483	35	3	6	0	3	3	3	-2	2	0	0	0	-1	-2	-1	0	0	0	0	1	1	0	0	0	0
ρ_7	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	γ_+^-	γ_-^-
$\overline{\rho_7}$	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	γ_-^-	γ_+^-
ρ_8	990	-18	-10	0	3	6	2	-2	0	0	-1	α_+^-	α_-^-	0	0	0	0	1	α_+^-	α_-^-	0	0	α_+^-	α_-^-	1	1
$\overline{\rho_8}$	990	-18	-10	0	3	6	2	-2	0	0	-1	α_-^-	α_+^-	0	0	0	0	1	α_-^-	α_+^-	0	0	α_-^-	α_+^-	1	1
ρ_9	1035	27	35	0	6	3	-1	3	0	0	2	-1	-1	1	0	1	0	0	-1	-1	0	0	-1	-1	0	0
ρ_{10}	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2\alpha_+^-$	$2\alpha_-^-$	-1	0	1	0	-1	0	0	0	0	α_-^+	α_+^+	0	0
$\overline{\rho_{10}}$	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2\alpha_-^-$	$2\alpha_+^-$	-1	0	1	0	-1	0	0	0	0	α_+^+	α_-^+	0	0
ρ_{11}	1265	49	-15	5	8	-7	1	-3	0	1	0	-2	-2	1	0	0	-1	0	0	0	0	0	1	1	0	0
ρ_{12}	1771	-21	11	16	7	3	-5	-1	1	0	-1	0	0	-1	1	0	0	-1	0	0	1	1	0	0	0	0
ρ_{13}	2024	8	24	-1	8	8	0	0	-1	-1	0	1	1	0	-1	0	-1	0	1	1	-1	-1	1	1	0	0
ρ_{14}	2277	21	-19	0	6	-3	1	-3	-3	0	2	2	2	-1	1	0	0	0	0	0	0	0	-1	-1	0	0
ρ_{15}	3312	48	16	0	-6	0	0	0	-3	0	-2	1	1	0	1	1	0	0	-1	-1	0	0	1	1	0	0
ρ_{16}	3520	64	0	10	-8	0	0	0	0	-2	0	-1	-1	0	0	0	0	0	1	1	0	0	-1	-1	1	1
ρ_{17}	5313	49	9	-15	0	1	-3	-3	3	1	0	0	0	-1	-1	0	1	0	0	0	0	0	0	0	0	0
ρ_{18}	5544	-56	24	9	0	-8	0	0	-1	1	0	0	0	0	-1	0	1	0	0	0	-1	-1	0	0	1	1
ρ_{19}	5796	-28	36	-9	0	-4	4	0	1	-1	0	0	0	0	1	-1	-1	0	0	0	1	1	0	0	0	0
ρ_{20}	10395	-21	-45	0	0	3	-1	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	-1	-1

The *strong* conjecture says that these ϕ_g are the twisted elliptic genera of an $N = 4$ vertex operator superalgebra — we have nothing to say about this. One would have liked to prove Theorem A by explicitly constructing these M_{24}-representations H_n. This remains an important challenge. Instead, we prove Theorem A in two steps.

The first step in showing this is to prove that the H_n are *virtual characters*, i.e. linear combinations over the integers of irreducible characters. It can be easily shown that the coefficients $H_n(g)$ will be *integer-valued* class functions, but this only implies the class function is a linear combination over $\mathbb{Q}$ of irreducibles, with denominators dividing the order $|G|$ of the group (244 million, in our case!). We prove the H_n are virtual by verifying that for each n the quantities $H_n(g)$ satisfy certain congruences; we can verify this simultaneously for all n by studying mod p^α reductions of associated modular forms. (Our result Lemma 3 on mod n reductions of modular forms is a refinement of older results in the literature.)

The other and more difficult step in proving Theorem A, is to show that each $H_n(g)$ is a *nonnegative* (real) linear combination of irreducible characters. It is easy to reduce this to showing that for any $g \neq 1$, $|H_n(g)|$ is small compared to $H_n(1)$ for all $n > N$ for some N (the coefficients for $n \leq N$ are then checked explicitly). Thanks to [6] we know

$$H_n(g) = \mathcal{O}\left(\frac{e^{\pi\sqrt{8n-1}/2|g|}}{\sqrt{8n-1}}\right) \tag{1.1}$$

where $|g|$ denotes the order of g, so for all $n > N$, where N is sufficiently large, $H_n(1)$

will dwarf $H_n(g)$ for $g \neq 1$, but we need to make this *effective* by finding that N, and this is the hard part of this positivity argument. It is tempting to guess we could follow Rademacher's calculation of effective bounds for the partition numbers, which looks like it should be similar, but Rademacher's series was absolutely convergent whereas those in [6] aren't, so the argument is much more delicate. But the analogy with partition numbers is still useful: *generalised* Kloosterman sums arise here, and the corresponding zeta functions are what we need to bound; but Whiteman [51] showed long ago that the classical Kloosterman sum has an elegant expression as a sparse sum, and Hooley [37] explained how to use the theory of binary quadratic forms to bound series associated to similar sums, so we follow their lead. This involves though finding significant generalisations of Whiteman's and Hooley's results. As an aside, our inequalities yield an independent proof of the convergence of the Rademacher sum expressions in [6].

Integrality of these multiplicities is crucial. In contrast, the significance of positivity is not as clear to this author — after all, elliptic genus is a *signed* trace. What happens here should be contrasted with the $N = 4$ $c = 6$ toroidal theories, where the elliptic genus vanishes. Positivity could be a consequence of the minimality of this string theory (see e.g. the discussions in [44],[52]) — we return to this important question in the conclusion. In any case, it is intriguing to note that the elliptic genera of the M_{24}-twisted modules (these have been recently obtained in [31]) also appear to be true representations (of the appropriate central extension of centralisers in M_{24}).

Our methods of proving weak Moonshine are quite robust. In particular, they can surely be applied to the twisted twining Mathieu Moonshine elliptic genera of [31], the 'Umbral Moonshine' of [7] and the $\mathrm{PSL}_2(\mathbb{Z}_{11})$ $N = 2$ Moonshine of [21] (as well of course to Monstrous Moonshine itself).

It is a consequence of our proof that the multiplicities $\mathrm{mult}_{\rho_i}(H_n)$ tend to infinity with n, for each i, and all $\mathrm{mult}_{\rho_i}(H_n)$ are strictly positive for $n \geq 25$. See Theorem 4 below.

It is common in the Mathieu Moonshine literature to emphasise the presence of *mock* modular forms. Indeed, the sums $q^{-1/8}\sum_{n=0}^{\infty} H_n(g)q^n$ are mock modular for each g, not (usually) modular. We had no use however of mock modularity in this paper; however the true modularity of the derived functions we call f_g plays a crucial role.

In Section 3 we also obtain an evenness result conjectured by several authors:

Theorem B. *Each head character H_n is a sum of*

$$\{2, 2\rho_1, \rho_2 + \overline{\rho_2}, \rho_3 + \overline{\rho_3}, 2\rho_4, 2\rho_5, 2\rho_6, \rho_7 + \overline{\rho_7}, \rho_8 + \overline{\rho_8}, 2\rho_9, \\ \rho_{10} + \overline{\rho_{10}}, 2\rho_{11}, 2\rho_{12}, 2\rho_{13}, 2\rho_{14}, 2\rho_{15}, 2\rho_{16}, 2\rho_{17}, 2\rho_{18}, 2\rho_{19}, 2\rho_{20}\}$$

Theorems A and B are both used in [12] to prove a conjecture in Umbral Moonshine [7]. Indeed, their Theorem 1.2 is far stronger than Conjecture 5.11 in [7] (specialised to M_{24}), and a little weaker than Conjecture 5.12 in [7]. Incidentally, it is curious that imaginary quadratic fields play a role in Umbral Moonshine (see Section 5.4 of [7]) while the equivalent theory of positive-definite quadratic forms plays a crucial role in our Section 4 proof. Perhaps this can supply a deeper explanation for Conjecture 5.12 than would arise from arguments as in [12].

Theorems A and B also imply that the elliptic genus of Enriques surfaces can be decomposed into true characters of M_{12} [22] (their elliptic genus is half that of K3 surfaces, so positivity comes from Theorem A and integrality from Theorem B). The VOA or string theoretic interpretation of this observation seems quite obscure.

[30] made the intriguing proposal that in fact the much larger Conway groups Co_0 or Co_1 may act. We prove (Gerald Höhn notified us that he also has a proof):

Theorem C. *All H_n, as well as H_{00}, are restrictions of virtual representations of Co_0.*

This was a nontrivial obstruction to their suggestion. However, as we explain below, the virtual Co_0-representations are much larger than the underlying M_{24}-representations (i.e. in the restriction a large representation with negative multiplicity cancels a large representation with positive multiplicity, leaving the small remainder H_n). Indeed, it now seems few people expect Mathieu Moonshine to extend to Co_0. Of course Theorem C implies that the same conclusion necessarily holds when Co_0 is replaced there with any subgroup between M_{24} and Co_0. In the conclusion, we mention the other group that should be considered a candidate for enhanced symmetry in Mathieu Moonshine.

In our paper, and indeed in much work on the subject, no relation involving K3 is used or obtained. It seems possible to us that the connection with K3 may be illusory. This would be very disappointing. We return to this a little more in the concluding section.

Because a fairly wide range of readers are potentially interested in parts of this paper, and may wish to extend it to other Moonshines, we have endeavoured to keep it as accessible and as detailed as possible.

2 K3-Mathieu Moonshine: Review

Throughout this paper, write $e(x) = \exp(2\pi \mathrm{i} x)$. For readability we will often collapse $a \equiv b \pmod{m}$ to $a \equiv_m b$. We will write $|g|$ for the order $|\langle g \rangle|$ of the element g of a group. An excellent review of superconformal field theories and their moduli spaces and elliptic genera, including $N = 4$ $c = 6$, is included in [52].

The elliptic genus of a sigma model with Calabi-Yau target is defined to be [54]

$$\phi(\tau, z) = \mathrm{Tr}_{H_{RR}}(q^{L_0 - c/24} e(zJ_0)(-1)^F \overline{q}^{\overline{L}_0 - c/24}), \tag{2.1}$$

where $q = e(\tau)$. Here, L_0 resp. $\overline{L}_0$ are the holomorphic resp. antiholomorphic Virasoro generators defining the grading, J_0 is an operator of $N = 2$ supersymmetry defining the charge, and F is the Fermionic number operator. The central charge c equals $3d$, where d is the (complex) dimension of the Calabi-Yau target. $\phi(\tau, z)$ will be independent of $\overline{q}$, by virtue of the supersymmetry in the Ramond sector, and depends only on the topological type (rather than the complex structure) of the target. It is a *weak Jacobi form* of index $m = d/2$ and weight $k = 0$ for $\mathbb{Z}^2 \rtimes \mathrm{SL}_2(\mathbb{Z})$. This means that ϕ is a holomorphic function of $z \in \mathbb{C}$ and $\tau \in \mathbb{H}$, the upper half-plane, which is modular with respect to τ and quasi-periodic with respect to z: for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $l, l' \in \mathbb{Z}$,

$$\phi\left(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d}\right) = (c\tau + d)^k e(mcz^2/(c\tau + d))\phi(\tau, z) \ , \tag{2.2}$$

$$\phi(\tau, z + l\tau + l') = e(-\mathrm{i}m\,(l^2\tau + 2lz))\phi(\tau, z) \ , \tag{2.3}$$

and in addition has Fourier expansion

$$\phi(\tau, z) = \sum_{n \geq 0, l \in \mathbb{Z}} c(n, l) q^n y^l \ , \tag{2.4}$$

for $y = e(z)$. Consistency of these conditions requires $c(n, l) = (-1)^k c(n, -l)$. The prefix 'weak' denotes that the sum over n is allowed to start at 0, i.e. ϕ is merely holomorphic at the cusp $\tau = \mathrm{i}\infty$; for historical reasons to be a true *Jacobi form* requires a sum over $l^2 \leq 4mn$.

The theory of Jacobi forms is developed in [25]. The definition extends trivially to weak Jacobi forms for other $\mathbb{Z}^2 \rtimes \Gamma$, where Γ is any subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of finite index. Note that because we have insisted here on the full group $\mathbb{Z}^2$ of translations, Γ must be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ because it must act on that group of translations. This is significant because it means we will lose the genus-0 property which played so important a role in Monstrous Moonshine — recall that many of the Monstrous Moonshine Fuchsian groups aren't subgroups of $\mathrm{SL}_2(\mathbb{Z})$.

Nevertheless, the genus-0 property is what made Monstrous Moonshine special (and rather mysterious), so we should seek an analogue for it in Mathieu Moonshine. [6] have an intriguing proposal: that the mock modular form $q^{-1/8} \sum_n H_n(g) q^n$ equals a certain regularised Rademacher sum. The relation of such a property (in weight-0) to the genus-0 property was established in [17].

The algebra of weak Jacobi forms for $\mathbb{Z}^2 \rtimes \Gamma$ of even weight and integral index, is the polynomial algebra $\mathfrak{m}_\Gamma[\phi_{0,1}, \phi_{-2,1}]$, where $\mathfrak{m}_\Gamma$ is the ring of holomorphic modular forms for Γ, and

$$\phi_{0,1}(\tau, z) = 4\left(\frac{\theta_2(\tau, z)^2}{\theta_2(\tau, 0)^2} + \frac{\theta_3(\tau, z)^2}{\theta_3(\tau, 0)^2} + \frac{\theta_4(\tau, z)^2}{\theta_4(\tau, 0)^2}\right) \ , \ \phi_{-2,1}(\tau, z) = -\frac{\theta_1(\tau, z)^2}{\eta(\tau)^6} \ , \tag{2.5}$$

for the classical Jacobi theta series θ_i and Dedekind eta function η. (For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ this is Proposition 9.3 in [25], but the proof for arbitrary finite-index Γ is the same.) In particular, every weight-0 index-1 weak Jacobi form $F(\tau, z)$ for $\mathbb{Z}^2 \rtimes \Gamma$ can be expressed as $F = h\phi_{0,1} + f\phi_{-2,1}$, where $h \in \mathbb{C}$ and f is a holomorphic weight-2 modular form for Γ. For nonlinear σ-models on K3, this implies the elliptic genus $\phi(\tau, z)$ must be proportional to $\phi_{0,1}$ and therefore will be constant throughout the 40-dimensional moduli space of K3 sigma models, something we already knew.

The worldsheet description is of an $N = 2$ vertex operator super algebra (VOSA). The holonomy of K3 surfaces extends the $N = 2$ to $N = 4$. Therefore, the space H_{RR} in fact carries a representation of the $N = 4$ superconformal algebra at $c = 6$, so

$$\phi_{K3}(\tau, z) = A_{00} ch^s_{1/4,0}(\tau, z) + \sum_{n=0}^{\infty} A_n ch^l_{n+1/4,1/2}(\tau, z)\,, \tag{2.6}$$

where $ch^*_{h,j}(\tau, z)$ is the elliptic genus of the appropriate level 1 $N = 4$ representation ($s =$short=massless=BPS, $l =$long=massive=non-BPS) and $ch^l_{1/4,1/2}$ is to be interpreted as $ch^s_{0,0} + 2ch^s_{1/4,1/2}$. Eguchi, Ooguri & Tachikawa [24] remarked that these coefficients A_n, previously calculated in [23], appear to be dimensions of M_{24}-representations (except for $n = 0$). More precisely, $A_{00} = 20$ is the dimension of the virtual representation $H_{00} = \rho_1 - 3\rho_0 = \rho_1 - 3$, and A_n is the dimension of H_n, where H_0 is the virtual representation -2, and the next few are

$$\begin{aligned}
&H_1 = \rho_2 + \overline{\rho_2}\,, \quad H_2 = \rho_3 + \overline{\rho_3}\,, \quad H_3 = \rho_7 + \overline{\rho_7}\,, \quad H_4 = 2\rho_{14}\,, \quad H_5 = 2\rho_{19}\,,\\
&H_6 = 2\rho_{20} + 2\rho_{16}\,, \quad H_7 = 2\rho_{20} + 2\rho_{19} + 2\rho_{18} + 2\rho_{17} + 2\rho_{13} + 2\rho_{12}\,,\\
&H_8 = 6\rho_{20} + 2\rho_{19} + 2\rho_{18} + 4\rho_{17} + 2\rho_{16} + 2\rho_{15} + 2\rho_{14} + 2\rho_{11} + \rho_{10} + \overline{\rho_{10}} + \rho_8 + \overline{\rho_8}\,,\\
&H_9 = 10\rho_{20} + 8\rho_{19} + 8\rho_{18} + 4\rho_{17} + 4\rho_{16} + 4\rho_{15} + 2\rho_{14} + 2\rho_{13} + 2\rho_{12} + 2\rho_{10} + 2\overline{\rho_{10}}\\
&\qquad\quad +2\rho_9 + 2\rho_7 + 2\overline{\rho_7} + 2\rho_6\,.
\end{aligned}$$

(Incidentally, [21] note that expanding the elliptic genus into $N = 2$ characters (rather than $N = 4$) seems to carry a representation of $\mathrm{PSL}_2(\mathbb{Z}_{11})$, a simple group of order 660.)

This observation of [24] is very reminiscent of the Monstrous Moonshine observation of McKay who noted that the expansion coefficients of the Hauptmodul J-function seem to be dimensions of Monster group $\mathbb{M}$ representations V_n (see e.g. [32] for a review and references). Thompson suggested considering what are now called the *McKay-Thompson series*, defined formally for all $g \in \mathbb{M}$ by

$$T_g(\tau) = \sum_{n=-1}^{\infty} \mathrm{ch}_{Vn}(g)\, q^n\,, \tag{2.7}$$

where $q = e^{2\pi \mathrm{i}\tau}$. Note that these T_g are constant on each of the 194 conjugacy classes of $\mathbb{M}$ — they are called *class functions.* After staring at their first few coefficients, Conway & Norton [9] conjectured that these T_g are modular functions (in fact Hauptmoduls=normalised generators of the field of modular functions) for some genus-0 Fuchsian group Γ_g commensurable with (but not in general a subgroup of) $\mathrm{SL}_2(\mathbb{Z})$. To make this conjecture more precise, we should turn the logic around and associate to each conjugacy class K_g in $\mathbb{M}$ a uniquely specified modular function T_g. From their Fourier coefficients we obtain a sequence χ_n of *class functions*, i.e. linear combinations over $\mathbb{C}$ of irreducible characters of $\mathbb{M}$. Because the coefficients of the T_g are rational (in fact integral), it is immediate that this linear combination can be taken over $\mathbb{Q}$, but it is very hard to show they can be taken over $\mathbb{Z}$ (i.e. that these class functions χ_n are virtual characters) and even harder to show that they are combinations over $\mathbb{Z}_{\geq 0}$ (i.e. that the χ_n are actually characters of $\mathbb{M}$-representations M_n). Atkin, Fong & Smith (see [48]) tried to prove with the aid of a computer that these χ_n are indeed true characters, and managed to reduce the proof to a fairly plausible statement they called Conjecture 2.3 (it is often claimed in the literature — see e.g. [32] — that [48] proves the χ_n are true characters, but this is false). That Conway-Norton conjecture was finally proved in [2], independently of [48]; thanks to this work, we now know much more: there is a vertex operator algebra (VOA) $\mathcal{V}^\natural$ (constructed in [27]) with automorphism group $\mathbb{M}$, whose *twining characters* $\mathrm{Tr}_{\mathcal{V}^\natural} g q^{L_0 - 1}$ equal the McKay–Thompson series T_g (the '-1' in the exponent is the usual shift by $c/24$). In other words, the $\mathbb{M}$-modules V_n are the eigenspaces of $V^\natural$ with respect to the Virasoro operator L_0. (The adjective 'twining' is short for 'intertwining', and is used as an alternative to the over-used word 'twisted'.)

Thus we are led to test further this Mathieu Moonshine observation of [24], by introducing for each $g \in M_{24}$

$$\phi_g(\tau, z) = \mathrm{ch}_{H00}(g)\, \mathrm{ch}_{1/4,0}(\tau, z) + \sum_{n=0}^{\infty} \mathrm{ch}_{Hn}(g)\, \mathrm{ch}_{n+1/4,1/2}(\tau, z)\,. \tag{2.8}$$

These should be the *twining elliptic genera*

$$\phi_g(\tau, z) = \mathrm{Tr}_{H_{RR}} g q^{L_0 - c/24} e(zJ_0)(-1)^F \overline{q}^{\overline{L}_0 - c/24}\,, \tag{2.9}$$

although that expression is purely formal as none of these K3 sigma models will have a symmetry consisting of all of M_{24}. In fact [42], the group of symplectic automorphisms of any K3 surface will typically be trivial, will never exceed order 960, and will only contain elements of order ≤ 8. And no automorphism, symplectic or otherwise, of a K3 surface can have order 23. But ignore these subtleties for now. Twining elliptic genera should have good modular properties — we would expect them to be weight-0 index-1 Jacobi forms for $\mathbb{Z}^2 \rtimes \Gamma_0(|g|)$ (recall we write $|g|$ for the order of g) though possibly with multiplier. This is enough to guess with effort these ϕ_g from the first few coefficients (just as was done in Monstrous Moonshine by Conway & Norton).

Conjectured expressions for ϕ_g, for all $g \in M_{24}$ (constant on conjugacy classes of course) were obtained in [5, 28, 29, 20]. We have

$$\phi_g(\tau, z) = \frac{w_g}{12} \phi_{0,1}(\tau, z) + f_g(\tau)\, \phi_{-2,1}(\tau, z)\,, \tag{2.10}$$

where $w_g \in \mathbb{Z}$ is the *Witten index* given in Table 2, and $f_g(\tau)$ is some holomorphic weight-2 modular form (with trivial multiplier) for $\Gamma_0(|g|h_g)$ for h_g in Table 2. w_g equals the character value of $\rho_1 + 1$ at g; note that it vanishes iff K_g doesn't intersect M_{23} — this isn't deep: $\rho_1 + 1$ is the permutation representation, so K_g intersects M_{23} iff $\rho_1 + 1$ has a fixed point, iff its character value (which equals the number of fixed-points) is nonzero. h_g is the length of the shortest cycle in the cycle shape of g. Note that in all cases, $h_g | \gcd(N, 12)$ (where N is the order of the element), and

$h_g = 1$ iff the conjugacy class K_g intersects M_{23}. We include for later use the order of the centraliser $C_{M24}(g)$ of g in M_{24}, and the index of $\Gamma_0(|g|)$ in $\mathrm{SL}_2(\mathbb{Z})$ (the index of $\Gamma_0(\prod_i p_i^{a_i})$, where p_i are distinct primes, is $\prod_i (p_i+1)p_i^{a_i-1}$), as well as the bound $a \times 10^b$ obtained by Theorem 3 below. In Table 2 and elsewhere, we write $7AB$ for $7A \cup 7B$ etc; we collect together conjugacy classes like this when the corresponding Jacobi forms ϕ_g coincide.

Table 2. Data for M_{24} conjugacy classes

K_g	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7AB	8A	10A	11A	12A	12B	14AB	15AB	21AB	23AB
h_g	1	1	2	1	3	2	1	4	1	1	6	1	1	2	1	2	12	1	1	3	1
w_g	24	8	0	6	0	0	4	0	4	2	0	3	2	0	2	0	0	1	1	0	1
$\lvert C(g)\rvert$	$\lvert M_{24}\rvert$	21504	7680	1080	504	384	128	96	60	24	24	42	16	20	11	12	12	14	15	21	23
index	1	3	3	4	4	6	6	6	6	12	12	8	12	18	12	24	24	24	24	32	24
$a\times 10^b$	3 6	5.7 6	2.1 9	4.6 7	2.3 12	2.9 11	2.1 8	5.1 14	6.8 8	2.9 9	1.5 18	7.5 9	2.2 10	3.5 14	2.6 11	1.5 15	5.8 23	1.7 12	2.9 12	8.7 18	8.2 13

Extracting f_g from (2.10), we obtain

$$\begin{aligned} f_g(\tau) &= w_g \frac{\theta_3^4+\theta_4^4}{12} - w_g\,\theta_3\theta_4\left(\frac{1}{4} + \sum_{n=1}^{\infty}\frac{q^{n(n+1)/2}}{1+q^n}\right) - \eta^3 q^{-1/8}\sum_{n=0}^{\infty} H_n(g) q^n \\ &= -\eta^3 q^{-1/8} \sum_{n=0}^{\infty}\Big(H_n(g) - \frac{w_g}{24}H_n(1)\Big)\, q^n\,. \end{aligned} \tag{2.11}$$

These modular forms $f_g(\tau)$ are given explicitly (i.e. independently of knowing the $H_n(g)$) for all $g \in M_{24}$ in [30, 20] (see also Table 2 of [6]). Indeed, recall the Eisenstein series $E_2(\tau) = 1 - 24q - 72q^2 - \cdots$, a quasi-modular form for $\mathrm{SL}_2(\mathbb{Z})$. Then for each integer $n > 1$,

$$E_2^{(n)}(\tau) := \frac{1}{n-1}(E_2(n\tau) - E_2(\tau)) = 1 + \frac{24}{n-1}\sum_{k=1}^{\infty}\sigma_1(k)\left(q^k - q^{nk}\right) \tag{2.12}$$

is a holomorphic modular form of weight 2 for $\Gamma_0(n)$ with trivial multiplier ($\sigma_1(k) = \sum_{d|k} d$). Writing $\eta(n) := \eta(n\tau)$ and $\eta = \eta(\tau)$, we have: $f_{1A} = 0$,

$$\begin{aligned} f_{2A} &= \frac{4}{3}E_2^{(2)} = \frac{4}{3} + 32q + 32q^2 + 128q^3 + 32q^4 + 192q^5 + 128q^6 + 256q^7 + 32q^8 + 416q^9 \cdots, \\ f_{2B} &= 2\frac{\eta^8}{\eta(2)^4} = 2 - 16q + 48q^2 - 64q^3 + 48q^4 - 96q^5 + 192q^6 - 128q^7 + 48q^8 - 208q^9 + \cdots, \\ f_{3A} &= \frac{3}{2}E_2^{(3)} = \frac{3}{2} + 18q + 54q^2 + 18q^3 + 126q^4 + 108q^5 + 54q^6 + 144q^7 + 270q^8 + 18q^9 + \cdots, \\ f_{3B} &= 2\frac{\eta^6}{\eta(3)^2} = 2 - 12q + 18q^2 + 24q^3 - 84q^4 + 36q^5 + 72q^6 - 96q^7 + 90q^8 + 24q^9 + \cdots, \\ f_{4A} &= 2\frac{\eta(2)^8}{\eta(4)^4} = 2 - 16q^2 + 48q^4 - 64q^6 + 48q^8 + 0q^9 + \cdots, \\ f_{4B} &= 2E_2^{(4)} - \frac{1}{3}E_2^{(2)} = \frac{5}{3} + 8q + 40q^2 + 32q^3 + 40q^4 + 48q^5 + 160q^6 + 64q^7 + 40q^8 + 104q^9 + \cdots, \\ f_{4C} &= 2\frac{\eta^4\eta(2)^2}{\eta(4)^2} = 2 - 8q + 32q^3 - 16q^4 - 48q^5 + 64q^7 + 48q^8 - 104q^9 + \cdots, \end{aligned}$$

$$\begin{aligned}
f_{5A} &= \frac{5}{3}E_2^{(5)} = \frac{5}{3} + 10q + 30q^2 + 40q^3 + 70q^4 + 10q^5 + 120q^6 + 80q^7 + 150q^8 + 130q^9 + \cdots\,, \\
f_{6A} &= \frac{5}{2}E_2^{(6)} - \frac{1}{2}E_2^{(3)} - \frac{1}{6}E_2^{(2)} = \frac{11}{6} + 2q + 14q^2 + 26q^3 + 38q^4 + 12q^5 + 38q^6 + 16q^7 + 86q^8 + 98q^9 + \cdots\,, \\
f_{6B} &= 2\frac{\eta^2\eta(2)^2\eta(3)^2}{\eta(6)^2} = 2 - 4q - 6q^2 + 8q^3 + 12q^4 + 12q^5 - 24q^6 - 32q^7 - 6q^8 + 8q^9 + \cdots\,, \\
f_{7AB} &= \frac{7}{4}E_2^{(7)} = \frac{7}{4} + 7q + 21q^2 + 28q^3 + 49q^4 + 42q^5 + 84q^6 + 7q^7 + 105q^8 + 91q^9 + \cdots\,, \\
f_{8A} &= \frac{7}{3}E_2^{(8)} - \frac{1}{2}E_2^{(4)} = \frac{11}{6} + 4q + 12q^2 + 16q^3 + 44q^4 + 24q^5 + 48q^6 + 32q^7 + 44q^8 + 52q^9 + \cdots\,, \\
f_{10A} &= 2\frac{\eta^3\eta(2)\eta(5)}{\eta(10)} = 2 - 6q - 2q^2 + 16q^3 - 2q^4 - 6q^5 - 8q^6 - 8q^7 - 2q^8 + 2q^9 + \cdots\,, \\
f_{11A} &= \frac{11}{6}E_2^{(11)} - \frac{22}{5}\eta^2\eta(11)^2 = \frac{11}{6} + 22q^2 + 22q^3 + 22q^4 + 22q^5 + 44q^6 + 44q^7 + 66q^8 + 66q^9 + \cdots\,, \\
f_{12A} &= 2\frac{\eta^3\eta(4)^2\eta(6)^3}{\eta(2)\eta(3)\eta(12)^2} = 2 - 6q + 2q^2 + 6q^3 - 6q^4 + 12q^5 - 10q^6 - 6q^8 - 6q^9 + \cdots\,, \\
f_{12B} &= 2\frac{\eta^4\eta(4)\eta(6)}{\eta(2)\eta(12)} = 2 - 8q + 6q^2 + 8q^3 - 4q^4 - 24q^6 + 16q^7 - 6q^8 + 16q^9 + \cdots\,, \\
f_{14AB} &= \frac{91}{36}E_2^{(14)} - \frac{7}{12}E_2^{(7)} - \frac{1}{36}E_2^{(2)} - \frac{14}{3}\eta\eta(2)\eta(7)\eta(14) \\
&= \frac{23}{12} - 3q + 11q^2 + 16q^3 + 11q^4 + 10q^5 + 16q^6 + 25q^7 + 39q^8 + 17q^9 + \cdots\,, \\
f_{15AB} &= \frac{35}{16}E_2^{(15)} - \frac{5}{24}E_2^{(5)} - \frac{1}{16}E_2^{(3)} - \frac{15}{4}\eta\eta(3)\eta(5)\eta(15) \\
&= \frac{23}{12} - 2q + 9q^2 + 13q^3 + 16q^4 + 13q^5 + 24q^6 + 14q^7 + 15q^8 + 28q^9 + \cdots\,, \\
f_{21AB} &= \frac{7}{3}\frac{\eta^3\eta(7)^3}{\eta(3)\eta(21)} - \frac{1}{3}\frac{\eta^6}{\eta(3)^2} = 2 - 5q - 3q^2 + 10q^3 + 7q^4 - 6q^5 - 12q^6 - 5q^7 + 6q^8 + 10q^9 + \cdots\,, \\
f_{23AB} &= \frac{23}{12}E_2^{(23)} - 69\eta^2\eta(23)^2 + 92\eta(2)^2\eta(46)^2 + 92\eta\eta(2)\eta(23)\eta(46) + 23\frac{\eta^3\eta(23)^3}{\eta(2)\eta(46)} \\
&= \frac{23}{12} + 23q^3 + 23q^4 + 23q^6 + 46q^8 + 23q^9 + \cdots\,.
\end{aligned}$$

Equation (7.16) of [13] tells us

$$\eta^3 q^{-1/8}\sum_{n=0}^{\infty} H_n(1)\,q^n = 48F^{(2)}(\tau) - 2E_2(\tau)\,, \tag{2.13}$$

where $F_2^{(2)}(\tau) = \sum_{n>m>0,n\not\equiv_2 m}(-1)^n m q^{mn/2}$. We read off from (2.12) and (2.13) that

$$E_2^{(2)} \equiv 1 \pmod{24}\,, \tag{2.14}$$

$$E_2^{(3)} \equiv 1 \pmod{12}\,, \tag{2.15}$$

$$\eta^3 q^{-1/8}\sum_{n=0}^{\infty} H_n(1)\,q^n \equiv -2 \qquad \pmod{48}\,. \tag{2.16}$$

These will be used in the Section 3 proofs. (We thank Thomas Creutzig and Gerald Höhn for bringing (2.13) and its consequence (2.16) to our attention.)

Lemma 1. [29] *Let* $g \in M_{24}$ *have order* $|g|$, *with parameter* h *given in Table 2. Then the twining character* $\phi_g(\tau, z)$ *is a Jacobi form of index 1 and weight 0 under* $\Gamma_0(|g|)$ *with multiplier defined by*

$$\phi_g\left(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\right) = e^{2\pi i cd/(|g|\,h)} e^{2\pi i cz^2/(c\tau+d)} \phi_g(\tau, z)\,. \tag{2.17}$$

The function $f_g(\tau)$*is a holomorphic modular form of weight 2 for* $\Gamma_0(|g|)$ *with multiplier* $e^{2\pi i cd/(|g|\,h)}$, *i.e.* $f_g\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 e^{2\pi i cd/(|g|\,h)} f_g(\tau)\ \forall \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(|g|)$.

$\Gamma_0(n)$ as usual consists of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $n|c$. The appearance of both $\Gamma_0(|g|)$ and the multiplier in Lemma 1 are not at all mysterious, as explained in the concluding section.

Again, reversing the logic, we can use these conjectured expressions for ϕ_g to define class functions $H_{00}(g)$ and $H_n(g)$. It was checked by brute force that these class functions are in fact true characters, for all $n \leq 500$ (except H_0 and H_{00} which are merely virtual). This is our Theorem A.

One would certainly want more than weak Moonshine (Theorem A) — e.g. we should have an M_{24}-worth of twisted twining elliptic genera with nice modularity. This is now done [31]. Unlike Monstrous Moonshine, where the algebraic content comes from the VOA $V^\natural$, the algebraic (not to mention geometric and physical) meaning of K3-Mathieu Moonshine is still unclear. Although we have learned much in the two years or so since [24], we still don't really know the right questions to ask.

3 Weak K3-Mathieu Moonshine I: Integrality

In this section and the next, we prove Theorems A and B, which were stated in Section 2. Our proof falls into two independent steps: *integrality* (this section) and *positivity* (next section).

Let G be any finite group. Let $\mathcal{P}_G$ be the set of all pairs (p, K_g) where p is a prime dividing the order $|G|$, K_g is a *p-regular* conjugacy class in G (i.e. the order $|\langle g\rangle|$ of g is coprime to p), and p divides the order $|C_G(g)|$ of the centraliser. There are precisely 22 pairs in $\mathcal{P}_G$ for $G = M_{24}$, which we list in Table 3. For reasons to become clear shortly, we include the highest power p^π of p dividing $|C_G(g)|$, and the p'*-section* S which we define next paragraph.

Table 3. Data for virtual character proof

p	2	2	2	2	2	2	3	3	3	3	3	3	3	3	5	5	5	7	7	7	11	23
K_g	1A	3A	3B	5A	7A	7B	1A	2A	2B	4A	4C	5A	7A	7B	1A	2B	3A	1A	2A	3B	1A	1A
π	10	3	3	2	1	1	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
S	1A,2AB 4ABC 8A	3A 6A 12A	3B 6B 12B	5A 10A	7A 14A	7B 14B	1A 3A 3B	2A 6A	2B 6B	4A 12A	4C 12B	5A 15A 15B	7A 21A	7B 21B	1A 5A	2B 10A	3A 15A 15B	1A 7A 7B	2A 14A 14B	3B 21A 21B	1A 11A	1A 23A 23B
q^n	q^4	q^8	q^{48}	q^6	q^4	q^4	q^2	q^2	q^2	q^4	q^6	q^4	q^{16}	q^{16}	q^1	q^3	q^4	q^1	q^4	q^5	q^2	q^4

For any $h \in G$ of order $n = p^k m$, where $\gcd(p, m) = 1$, find $a, b \in \mathbb{Z}$ such that $1 = ap^k + bm$. Then $h = h_{p'} h_p$ where the *p'-part* $h_{p'} = h^{ap^k}$ has order m and the *p-part* $h_p = h^{bm}$ has order p^k. The p'-section S for the pair $(p, K_g) \in \mathcal{P}_G$ is defined to be $S = \{h \in G \,|\, h_{p'} \in K_g\}$. S is clearly the union of conjugacy classes; in Table 3 we list those conjugacy classes.

Suppose we have an integer-valued class function $c : G \to \mathbb{Z}$, i.e. c is constant on conjugacy classes. We are interested in H being one of the head characters H_n, and we want to prove it is the character of an M_{24}-representation. The hard part is to prove that A is a virtual character, i.e. a linear combination over $\mathbb{Z}$ of irreps. The starting point is the following characterisation of virtual characters, attributed to Thompson and based on the classical theorem of Brauer. It is a refinement of the following basic idea, which the reader can verify for himself: If χ is an integer-valued character and g is an element of order a prime power p^π, then $\chi(g) \equiv \chi(e) \pmod{p}$.

Choose any $(p, K_g) \in \mathcal{P}_G$ and write the p'-section S as a disjoint union $\cup_{i=1}^k K_i$ of conjugacy classes, for some k depending on (p, S). Let R_p be the tensor product $\widehat{\mathbb{Z}}_p \otimes_{\mathbb{Z}} \mathbb{Z}[\xi_{|G|}]$, where $\widehat{\mathbb{Z}}_p$ are the p-adic integers and $\xi_n = e^{2\pi \mathrm{i}/n}$. R_p is introduced to make formal sense of the orthogonality relations we're about to introduce; the ring $\mathbb{Z}[\xi_{|G|}]$ arises because all irreducible G-characters take values there. Define $\mathcal{M}_{(p,K_g)}$ to be the set of all k-tuples $(\ell_1, \ldots, \ell_k) \in R_p^k$ such that $\sum_{i=1}^k \ell_i \chi(K_i) \in p^\pi R_p$ for all irreducible characters χ. If c is a virtual character, then c is likewise orthogonal to $\mathcal{M}_{(p,S)}$. What is important to us is the converse:

Lemma 2. [48] *Let R_p and $\mathcal{M}_{(p,K_g)}$ be as above. Let $c : G \to \mathbb{Z}$ be an integer-valued class function of G. Then c is a virtual character of G, i.e. a linear combination over $\mathbb{Z}$ of irreducibles, if $\sum_{i=1}^k \ell_i c(K_i) \in p^\pi R_p$ for all $(\ell_1, \ldots, \ell_k) \in \mathcal{M}_{(p,K_g)}$ and all $(p, K_g) \in \mathcal{P}_G$.*

Of course we want to apply this to $c = H_n$, for each $n > 0$. The reason this characterisation is helpful is that much is known about reductions of modular forms modulo powers of primes, as we will see.

Suppose $f(\tau) = \sum_{n=0}^\infty f_n q^n, g(\tau) = \sum_{n=0}^\infty g_n q^n$ are holomorphic modular forms of weight $k \in \mathbb{Z}$ for some finite-index subgroup Γ and multiplier $\mu : \Gamma \to \mathbb{C}^\times$, and μ has finite order M. If the Fourier coefficients f_n, g_n are equal for all $n \leq k\|\mathrm{SL}_2(\mathbb{Z})/\Gamma\|/12$, then $f = g$. This is an immediate consequence of the classical valence formula for Γ applied to $(f - g)^M$. Incidentally, because of this and Lemma 1, the q-expansions provided earlier are more than enough to uniquely specify f_g (and hence ϕ_g) — in fact the coefficients $n \leq 5$ suffice. What is much more surprising is that this also applies to the modular forms mod prime powers:

Lemma 3. *Suppose $f(\tau) = \sum_{l \geq 0} f_l q^l \in \mathbb{Q}[[q^{1/N}]]$ and $g(\tau) = \sum_{l \geq 0} g_l q^l \in \mathbb{Q}[[q^{1/N}]]$ are holomorphic modular forms of rational weight $k \in \mathbb{Q}$ and with some multiplier μ for some subgroup Γ of $SL_2(\mathbb{Z})$. Let $m := \|\mathrm{SL}_2(\mathbb{Z})/\Gamma\|$ be the index. We require that μ has finite order, i.e. that all values $\mu(\gamma)$ for $\gamma \in \Gamma$ are Mth roots of 1 for some M.*

(a) *Suppose the Fourier coefficients f_l and g_l are integral for all l. Suppose we have an integer $n > 0$ such that $f_l \equiv g_l \pmod{n}$ for all $l \leq km/12$. Then $f_l \equiv g_l \pmod{p}$*

for all l.

(b) *Suppose now that the weight k is integral and that the kernel of μ is a congruence subgroup, i.e. contains $\Gamma(N')$ for some N'. If the coefficients f_l are integral for all $0 \leq l \leq km/12$, then $f_l \in \mathbb{Z}$ for all l.*

Proof. The key result upon which this Lemma rests is Sturm's Theorem [49], which is part (a) in the special case where the multiplier μ is trivial and the weight is integral and n is a prime (Sturm however allows the Fourier coefficients to be algebraic integers).

It suffices to prove the Lemma for $g = 0$. Start with part (a). Let p be any prime dividing n. Choose an integer $K > 0$ so that $Kk \in \mathbb{Z}$ and μ^K is identically 1. Then $F := f^K$ is a holomorphic modular form for Γ of integer weight Kk and trivial multiplier, whose Fourier coefficients F_l are multiples of the prime p for all $l \leq Kkm/12$ (since each F_l will be a sum of monomials in f_{l_i}'s at least one of which has $l_i \leq km/12$). Then Sturm's theorem applies, and we find that p divides *all* coefficients F_l. This must imply that p divides all coefficients of f (otherwise, choose the smallest l such that p doesn't divide f_l, and note that p would fail to divide F_{Kl}). This means that f/p will have integer coefficients, so f/p will obey all hypotheses of part (a) with now n replaced with n/p. Repeat with another prime dividing n/p, until you've reduced n to 1.

To see Lemma 3(b), let V be the space of all holomorphic weight k modular forms for $\Gamma(N')$ with trivial multiplier. V is finite-dimensional with an integral basis $f^{(i)} \in \mathbb{Z}[[q^{1/N'}]]$. Therefore we can write $f = \sum_i a_i f^{(i)}$ where $a_i \in \mathbb{Q}$. This means the coefficients of f have denominator bounded by the lcm of all denominators of the a_i. Let L be the smallest positive integer such that $Lf \in \mathbb{Z}[[q]]$. If p is a prime dividing L then p will divide the first $mk/12$ coefficients of Lf by the hypothesis on f, hence will divide all coefficients by Lemma 3(a), contradicting the minimality of L. This means $L = 1$, and we're done. **QED to Lemma 3**

Theorem 1. *Each H_n ($n \geq 1$) is a virtual character of M_{24} (i.e. a linear combination over $\mathbb{Z}$ of irreducible characters of M_{24}).*

Proof. First, we need to show the class functions H_n are integer-valued. We will do this by studying the f_g's. One complication is that the coefficients of $f_g(\tau)$ are not quite integers, as can be seen by the displayed values in Section 2. The problem is the constant term $\frac{1}{4}$ in the middle term of (2.11), and the 12 in the denominator of its first term. In fact only the constant term of $w_g\,(\theta_3^4 + \theta_4^4)/12 - w_g\theta_3\theta_4/4$ can fail to be integral. To see this, it suffices to show that $\theta_3\theta_4 \in 1 + 4q\mathbb{Z}[[q]]$ and $\theta_3^4 + \theta_4^4 \in 2 + 12q\mathbb{Z}[[q]]$. That these both hold modulo 4 follows from $\theta_3 = 1 + 2\sum + 2\sum'$ and $\theta_4 = 1 + 2\sum - 2\sum'$, using obvious notation. That $\theta_3^4 + \theta_4^4 \equiv 1$ (modulo 3) follows from comparing $(\theta_3^4 + \theta_4^4)^2$ and E_4 modulo 3, using Lemma 3 with $\Gamma = \Gamma(2)$.

From Lemma 3(b) and checking the integrality of the first few coefficients of f_g we learn now that all coefficients of $f_g - w_g\,(\theta_3^4 + \theta_4^4)/12 + w_g\theta_3^2\theta_4^2/4$ are integral (the theta function contribution kills the fractional part of the constant term). This means that all coefficients of each f_g are integral except possibly for the constant term, and also (from the first equality in (2.11)) that all coefficients of $\eta^3 q^{-1/8}\sum H_n(g)q^n$ are integral. Since $\eta^3 q^{-1/8}$ is invertible in the ring $\mathbb{Z}[[q]]$, we find that $H_n(g)\in\mathbb{Z}$ for all $n\geq 0$ and all $g\in M_{24}$.

This invertibility of $\eta^3 q^{-1/8}$ directly gives us the useful implication:

$$l f_g \in m\mathbb{Z}[[q]] \;\Rightarrow\; lH_n(g)\equiv l\frac{w_g}{24}H_n(1)\pmod{m} \tag{3.1}$$

for all $n\geq 0$, for any choice of $l,m\in\mathbb{Z}$ and all $g\in M_{24}$.

We need to verify, for each pair $(p,K_g)\in\mathcal{P}_{M24}$, that $\sum_{i=1}^{k}\ell_i H_n(K_i)\in p^\pi R_p$ for all $(\ell_1,\ldots,\ell_k)\in\mathcal{M}_{(p,K_g)}$.

A typical example is the pair $(3,2A)$. From Table 3 we see that $S=K_{2A}\cup K_{6A}$. Note that $\chi(2A)\equiv_3\chi(6A)$, for all irreducible χ, so we need to show $H_n(2A)\equiv_3 H_n(6A)$ for all n. We claim that $f_{2A}-4f_{6A}\in 3\mathbb{Z}[[q]]$. To see this, first note that $f_{2A}-4f_{6A}$ is a modular form for $\Gamma_0(6)$ with integer q-expansion and trivial multiplier, so by Lemma 3(a) it suffices to check that 3 divides the q^l-coefficient of $f_{2A}-4f_{6A}$ for all $0\leq l\leq 2$, which is trivial to do from the expansions collected in Section 2. By (3.1) or otherwise, this implies $H_n(2A)\equiv_3 H_n(6A)$ for all n and we're done. The pairs $(2,5A),(2,7AB),(3,2B),(3,4AC),(5,1A),(5,2B),(11,1A)$ are all handled similarly.

The pair $(3,5A)$ is also easy. Here $S=5A\cup 15A\cup 15B$, and we find $\mathcal{M}_{(3,5A)}=R_3$-span$\{(0,1,-1),(1,-1,0)\}$ using obvious notation. $H_n(15A)=H_n(15B)$ for all n follows from $f_{15A}=f_{15B}$; $H_n(5A)\equiv_5 H_n(15A)$ follows from $f_{5A}\equiv_3 4f_{15A}$, proved in the usual manner using $\Gamma_0(15)$. The pairs $(5,3A),(7,1A),(7,2A),(7,3B),(23,1A)$ are done similarly.

For $(3,7AB)$, we need $H_n(7AB)\equiv_3 H_n(21AB)$ for all n. Using (2.11), as well as the congruences (2.16),(2.15), this is equivalent to verifying that $4f_{7AB}-f_{21AB}\equiv_3 2E_2^{(3)}$, which by Lemma 3 for $\Gamma=\Gamma_0(63)$ requires checking up to q^{16}. For $(2,3A)$, $S=3A\cup 6A\cup 12A$ and $\mathcal{M}_{(2,3A)}=R_2$-span$\{(2,-2,0),(1,1,1)\}$, so we need to verify that $H_n(3A)\equiv_2 H_n(6A)$ and $H_n(3A)+H_n(6A)\equiv_8 2H_n(12A)$, for all n. Using now (2.14), this is equivalent to verifying that $f_{3A}\equiv_2 f_{6A}$ and $f_{3A}+9f_{6A}-2f_{12A}\equiv_8 -2E_2^{(2)}$. The former requires checking up to q^2 (use $\Gamma_0(6)$), while the latter requires checking up to q^8 (use $\Gamma_0(24)$). The pair $(2,3B)$ is handled similarly.

The pair $(3,1A)$ has $S=1A\cup 3A\cup 3B$ and $\mathcal{M}_{(3,1A)}=R_3$-span$\{(1,1,1),(0,9,3),(0,0,9)\}$, so we need to establish $H_n(1A)\equiv_3 H_n(3B)$ and $H_n(3A)\equiv_{27} 3H_n(3B)-2H_n(1A)$ for all n. These are equivalent to $f_{3B}\equiv_3 2E_2^{(3)}$ and $4f_{3A}-12f_{3B}\equiv_{27} 9E_2^{(3)}$.

Finally, $\mathcal{M}_{(2,1A)}$ is the span of $(1,1,1,1,1,1,1)$, $(-22,-6,2,2,-2,2,0)$, $(44,12,28,4,4,0,0)$, $(-24,8,24,8,0,0,0)$, $(-208,-16,16,0,0,0,0)$, $(448,64,0,0,0,0,0)$, using obvious notation. Therefore we need to verify

$$\begin{aligned}
&H_n(4C) \equiv H_n(8A) \pmod 2\,,\\
&H_n(4B) \equiv 2H_n(8A) - H_n(4C) \pmod 4\,,\\
&H_n(4A) \equiv H_n(4B) + 2H_n(4C) - 2H_n(8A) \pmod 8\,,\\
&H_n(2B) \equiv 3H_n(4A) + 4H_n(4B) + 2H_n(4C) + 8H_n(8A) \pmod{16}\,,\\
&H_n(2A) \equiv -H_n(2B) + 4H_n(4A) + 6H_n(4B) - 8H_n(8A) \pmod{64}\,,\\
&H_n(1A) \equiv 7H_n(2A) - 6H_n(2B) + 8H_n(4A) + 24H_n(4B) + 32H_n(4C) - 64H_n(8A) \pmod{1024}
\end{aligned}$$

for all n. These are equivalent to

$$\begin{aligned}
&2f_{4C} \equiv 2f_{8A} + E_2^{(2)} \pmod 4\,,\\
&f_{4B} \equiv 2f_{8A} - f_{4C} \pmod 4\,,\\
&f_{4A} \equiv f_{4B} + 2f_{4C} - 2f_{8A} \pmod 8\,,\\
&f_{2B} \equiv 3f_{4A} + 4f_{4B} + 2f_{4C} + 8f_{8A} + 8E_2^{(2)} \pmod{16}\,,\\
&f_{2A} \equiv -f_{2B} + 4f_{4A} + 6f_{4B} - 8f_{8A} \pmod{64}\,,\\
&7f_{2A} - 6f_{2B} + 8f_{4A} + 24f_{4B} + 32f_{4C} - 64f_{8A} \equiv 0 \pmod{1024}\,,
\end{aligned}$$

respectively, where all of these are modular forms for $\Gamma_0(16)$. *QED*

Refining this argument slightly gives us the evenness property stated in Theorem B. (Theorem B will follow from Theorem 2 and Theorem 3).

Theorem 2. *Each head character H_n is a linear combination over $\mathbb{Z}$ of*

$$\{2, 2\rho_1, \rho_2 + \overline{\rho_2}, \rho_3 + \overline{\rho_3}, 2\rho_4, 2\rho_5, 2\rho_6, \rho_7 + \overline{\rho_7}, \rho_8 + \overline{\rho_8}, 2\rho_9, \\ \rho_{10} + \overline{\rho_{10}}, 2\rho_{11}, 2\rho_{12}, 2\rho_{13}, 2\rho_{14}, 2\rho_{15}, 2\rho_{16}, 2\rho_{17}, 2\rho_{18}, 2\rho_{19}, 2\rho_{20}\}$$

Proof. Let χ be any M_{24}-class function, and $\rho \in \widehat{M_{24}}$. Write $\mathrm{mult}_\rho(\chi)$ for the multiplicity. Consider first ρ a *complex* character (i.e. $\overline{\rho} \ncong \rho$). We know that each $f_g(\tau) \in \mathbb{Q} + q\mathbb{Z}[[q]]$. The integrality of $H_n(g)$ for each n means that $\mathrm{mult}_\rho(H_n) = \mathrm{mult}_{\overline{\rho}}(H_n)$.

Much more difficult are the real M_{24}-irreps ρ (i.e. $\rho \cong \overline{\rho}$). Assume $\rho \neq 1, \rho_1$ for now. Define

$$M_\rho(\tau) := \sum_{n=0}^{\infty} \mathrm{mult}_\rho(H_n) q^n = -\sum_{K_g} |C_{M_{24}}(g)|^{-1} \overline{\rho(g)} f_g(\tau)\, q^{1/8} \eta(\tau)^{-3}\,, \tag{3.2}$$

using (2.11), where the second sum is over all conjugacy classes in M_{24}. We know from Theorem 1 that each $M_\rho(\tau) \in q\mathbb{Z}[[q]]$ (at least for $\rho \neq 1, \rho_1$). We want to show that in fact $M_\rho(\tau) \in 2q\mathbb{Z}[[q]]$. Of course we can ignore the constant terms of the f_g in the following.

First note from the M_{24} character table that $\rho(23A)$ and $\rho(23B)$ will be equal and integral. Also, the centraliser $C_{M24}(23AB)$ has odd order (namely, 23). Since

$f_{23AB}(\tau) \in \mathbb{Q}+q\mathbb{Z}[[q]]$, this means we can ignore the contribution of the classes $23AB$ to (3.2) — together they will contribute something in $\frac{2}{23}q\mathbb{Z}[[q]]$ and thus will not affect the value of $23M_i(\tau)$ modulo 2.

Identical arguments apply to the classes $21AB$ and $15AB$. The class $11A$ can likewise be ignored: its centraliser has odd order (namely 11), $\rho(11A) \in \mathbb{Z}$, and $f_{11A}(\tau) \in \mathbb{Q} + 2q\mathbb{Z}[[q]]$ (to see this, apply Lemma 3(a) to $\frac{1}{2}(f_{11A} + \frac{1}{6}E_2^{(11)})$). Finally, the classes $7AB$ and $14AB$ can all be dropped from (3.2): they all have the same order (namely 14) of centraliser, $\rho(7A) + \rho(7B) + \rho(14A) + \rho(14B) \in 4\mathbb{Z}$ (since $\overline{\rho} \cong \rho$), and the q^n-coefficients ($n > 0$) of f_{7AB} and f_{14AB} are integers congruent mod 2 (to see this, apply Lemma 3(a) to $3f_{14AB} - f_{7AB}$). Let $\widetilde{M_\rho}(\tau)$ be $23 \cdot 21 \cdot 5 \cdot 11$ times the sum in (3.2) restricted to the remaining classes $K_g \in \{1A, 2AB, 3AB, 4ABC, 5A, 6AB, 8A, 10A, 12AB\}$. Then $\widetilde{M_\rho}(\tau) \in q\mathbb{Z}[[q]]$ and $M_\rho \equiv_2 \widetilde{M_\rho}$ and we need to show that $\widetilde{M_\rho}(\tau) \in 2q\mathbb{Z}[[q]]$, or equivalently $\eta^3 q^{-1/8}\widetilde{M_\rho} \in 2q\mathbb{Z}[[q]]$.

Note that $\eta^3 q^{-1/8}\widetilde{M_\rho}$ is a linear combination (over $\mathbb{Q}$) of f_g with $\rho(g) \neq 0$ and $g \in \{1A, 2AB, 3AB, 4ABC, 5A, 6AB, 8A, 10A, 12AB\}$. Let N_ρ be the least common multiple of $h_g N_g$ over those g, where h_g is in Table 2 and N_g is the order of g. We know from Lemma 1 that f_g is a weight-2 modular form for $\Gamma_0(h_g N_g)$ with trivial multiplier, so $\eta^3 q^{-1/8}\widetilde{M_\rho}$ is a weight-2 modular form for $\Gamma_0(N_\rho)$ with trivial multiplier. We collect N_ρ in Table 4, together with the quantity $m_\rho/6$ where $m_\rho = \|\mathrm{SL}_2(\mathbb{Z})/\Gamma_0(N_\rho)\|$. By Lemma 3, we need to show that the first $m_\rho/6$ coefficients of $\eta^3 q^{-1/8}\widetilde{M_\rho}$ are even. As in (2.11) this is equivalent to showing that the first $m_\rho/6$ values of $\mathrm{mult}_\rho(H_n)$ are even. The evenness of these multiplicities $\mathrm{mult}_\rho(H_n)$ has been verified for all $n \leq 500$ and all real ρ, by Gaberdiel-Hohenegger-Volpato (private communication; see also [29]), which is far more than is necessary.

The proof for $\rho = 1$ and $\rho = \rho_1$ is similar, but (3.2) has to be modified. For those ρ define

$$\begin{aligned} F_\rho(\tau) &:= \sum_{K_g} |C_{M_{24}}(g)|^{-1}\overline{\rho(g)} f_g(\tau) + \frac{1}{12}E_2^{(2)}(\tau) \\ &= -q^{-1/8}\eta(\tau)^3 \sum_{n=0}^{\infty}\left(\mathrm{mult}_\rho(H_n) - \frac{1}{24}H_n(1)\right) q^n + \frac{1}{12}E_2^{(2)}(\tau) \\ &\equiv_2 -q^{-1/8}\eta(\tau)^3 \sum_{n=0}^{\infty}\mathrm{mult}_\rho(H_n)\, q^n\,, \end{aligned} \tag{3.3}$$

using (2.16), (2.14), (2.11) and the fact that $w_g = 1+\rho_1(g)$. Thus if we can show $F_\rho(\tau)$ has even coefficients, we will be done. The expression (3.3) together with Theorem 1 tells us F_ρ has integral Fourier coefficients. Define $\widetilde{F_\rho}$ as above by restricting the sum to classes $K_g \in \{1A, 2AB, 3AB, 4ABC, 5A, 6AB, 8A, 10A, 12AB\}$ and multiplying by $23 \cdot 21 \cdot 5 \cdot 11$; as above, it suffices to show the first 288 coefficients are even, equivalently that the multiplicities of both 1 and ρ_1 are even in all head characters

H_n for $n \leq 288$. This has been done in the aforementioned computer checks. **QED to Theorem 2**

Table 4. Data for Theorem B proof

ρ	1	ρ_1	ρ_4	ρ_5	ρ_6	ρ_9	ρ_{11}	ρ_{12}	ρ_{13}	ρ_{14}	ρ_{15}	ρ_{16}	ρ_{17}	ρ_{18}	ρ_{19}	ρ_{20}
N_ρ	$2^4 3^2 5$	$2^4 3^2 5$	$2^3 3\, 5$	$2^4 3^2 5$	$2^4 3\, 5$	$2^4 3^2$	$2^4 3^2$	$2^4 3^2 5$	$2^3 3^2 5$	$2^4 3^2 5$	$2^2 3^2 5$	$2\, 3^2$	$2^4 3\, 5$	$2^3 3\, 5$	$2^3 3\, 5$	2^4
$m_\rho/6$	288	288	48	288	96	48	48	288	144	288	72	6	96	48	48	48

4 Weak Mathieu Moonshine II: Positivity

In this section we prove that for each $n > 0$ and each irreducible $\rho \in \widehat{M_{24}}$, the multiplicities $\mathrm{mult}_{H_n}(\rho)$ are nonnegative. The difficult part of this positivity proof is effectively bounding a certain series (what we call $Z_{n;h}(3/4)$ below) which does not converge absolutely. There are (at least) two approaches for this: interpreting this as a Selberg-Kloosterman zeta function as in [47],[33], etc; or interpreting this as a sum over equivalence classes of quadratic forms as in Hooley [37]. The former method was used by [6] to prove convergence for the Rademacher sum expressions for the mock modular forms $q^{-1/8}\sum_{n=0}^{\infty} H_n(g)q^n$; the latter method (or rather its recent reincarnation [4]) was suggested in [19] as a way to prove convergence for $g = 1$. We need much more than convergence: we need an explicit bound, and for this we have found the second method more useful.

Let's begin with an elementary observation. By the triangle inequality,

$$\mathrm{mult}_{H_k}(\rho) = \sum_g \frac{H_k(g)}{|C_{M24}(g)|}\overline{\rho(g)} \geq \frac{H_k(1)}{|M24|}\rho(1) - \sum_{g\neq 1}\frac{|H_k(g)|}{|C_{M24}(g)|}|\rho(g)| \tag{4.1}$$

where the sum is over all conjugacy classes of M_{24}. Since there always is the trivial bound $\rho(1) \geq |\rho(g)|$ (which itself follows immediately from the triangle inequality), our strategy to show $\mathrm{mult}_{H_k}(\rho) > 0$ is to show $H_k(1)$ is much larger in modulus than the other character values $H_k(g)$, at least for k sufficiently large.

Choose any $g \in M_{24}$. Then from Lemma 1 we read that the multiplier of f_g is $\rho_{|g|;h_g}$, where $\rho_{n;h}$ sends $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \Gamma_0(n)$ to $e\left(\frac{cd}{nh}\right)$. The multiplier ϵ of η sends $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ (with $c > 0$) to

$$\frac{1-\mathrm{i}}{\sqrt{2}}\omega_{-d,c}\, e\left(\frac{a+d}{24c}\right), \tag{4.2}$$

where

$$\begin{aligned}\omega_{d,c} &= \prod_{\mu=1}^{k} \exp(\pi \mathrm{i} \left(\left(\frac{h\mu}{k}\right)\right)\left(\left(\frac{\mu}{k}\right)\right)) \\ &= \begin{cases} \left(\frac{-d}{c}\right) e\left(\frac{-1}{8}(c-1) + \frac{-1}{24}(c-\frac{1}{c})(2d+d'-d^2d')\right) & \text{if } c \text{ is odd} \\ \left(\frac{-c}{d}\right) e\left(\frac{-1}{8}(2-cd-d) + \frac{-1}{24}(c-\frac{1}{c})(2d+d'-d^2d')\right) & \text{if } c \text{ is even}\end{cases}.\end{aligned} \tag{4.3}$$

Here, $((x)) = x - \lfloor x \rfloor - 1/2$ unless $x \in \mathbb{Z}$ in which case $((x)) = 0$. We write here d' for any solution to $dd' \equiv_c 1$. $\left(\frac{d}{c}\right)$ denotes the Jacobi symbol. The first expression for $\omega_{d,c}$ is the most familiar; the second (due to Rademacher [45]) is more useful for us.

We learn in eq.(6.1) of [6] (see also [19] for the special case $g = 1$) that

$$H_k(g) = \frac{4\pi}{(8k-1)^{1/4}} \sum_{c=1}^{\infty} \frac{1}{|g|\, c} I_{1/2}\left(\frac{\pi}{2c\,|g|}\sqrt{8k-1}\right) S(k, |g|\, c, \epsilon^{-3}\rho_{|g|;h_g})\,, \tag{4.4}$$

where $I_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sinh(x)$ and S is (up to a constant) a generalised Kloosterman sum for $\Gamma_0(|g|)$:

$$S(k, nc, \epsilon^{-3}\rho_{n;h}) = \sum_{\substack{0<d\le nc\\ \gcd(d,nc)=1}} \omega_{d,nc}^{-3} e\left(\frac{-cd}{h}\right) e\left(\frac{kd}{nc}\right) \tag{4.5}$$

Fix integers $n, h, k > 0$ and define $L = \frac{\pi}{n}\sqrt{2k-1}$ (so $\pi\sqrt{2k-1}/(2nc) < 1$ if $c > L$). We have the elementary bounds:

$$\left| I_{1/2}(x) - \sqrt{\frac{2x}{\pi}} \right| \le \frac{1}{5}\sqrt{\frac{2x^5}{\pi}} \qquad \text{for } 0 < x < 1\,; \tag{4.6}$$

$$|I_{1/2}(x)| < \frac{e^x}{\sqrt{2\pi x}} \qquad \text{for all } 0 < x\,; \tag{4.7}$$

$$|S(k, nc, \epsilon^{-3}\rho_{n;h})| \le 1 \qquad \text{for all } c \in \mathbb{Z}_{>0} \tag{4.8}$$

(in fact Lemma 6 below will give us the Weil bound $|S| \le O(\sqrt{c})$, but (4.8) is adequate). These inequalities immediately imply the crude bounds

$$\begin{aligned}
&\left| \frac{1}{(8k-1)^{1/4}} \sum_{c=2}^{\lfloor L \rfloor} \frac{1}{c} I_{1/2}\left(\frac{\pi\sqrt{8k-1}}{2nc}\right) S(k, nc, \epsilon^{-3}\rho_{n;h}) \right| \\
&\qquad \le \frac{\sqrt{n}}{\pi\sqrt{8k-1}} \sum_{c=2}^{\lfloor L \rfloor} \frac{1}{\sqrt{c}} e^{\pi\sqrt{8k-1}/(2nc)} < \frac{1}{\sqrt{2n}} e^{\pi\sqrt{8k-1}/(4n)}\,,
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
&\left| \frac{1}{(8k-1)^{1/4}} \sum_{c=\lceil L \rceil}^{\infty} \frac{1}{nc} I_{1/2}\left(\frac{\pi\sqrt{8k-1}}{2nc}\right) S(k, nc, \epsilon^{-3}\rho_{n;h}) \right| \\
&\qquad \le \left| \sum_{c=\lceil L \rceil}^{\infty} (nc)^{-3/2} S(k, nc, \epsilon^{-3}\rho_{n;h}) \right| + \frac{\pi^2(8k-1)}{20n^{7/2}} \sum_{c=\lceil L \rceil}^{\infty} c^{-7/2} \\
&\qquad \le |Z_{n;h}(3/4)| + \frac{\pi\sqrt{8k-1}}{n^{5/2}} + \frac{7\pi^2(8k-1)}{80n^{7/2}}\,,
\end{aligned} \tag{4.10}$$

where $Z_{n;h}$ is the Selberg-Kloosterman zeta function

$$Z_{n;h}(s) = \sum_{c=1}^{\infty} S(k, nc, \epsilon^{-3}\rho_{n;h})\,(nc)^{-2s}\,. \tag{4.11}$$

Thus it suffices to find an effective bound for a Selberg-Kloosterman zeta function at $s = 3/4$. We know from [6] that the defining series (4.11) converges there, but it follows from Lemma 4 below that this convergence is not absolute, and in any event it seems difficult to use the analysis of [6] to obtain an explicit bound (as a function in k). The key first step in identifying such an effective bound is to rewrite these generalised Kloosterman sums more sparsely. Our calculation resembles that of [51]; a more elegant approach attributed to Selberg and worked out by Rademacher [46] is available but we were unable to generalise it to our context. Incidentally, Lemma 4 should imply that these S's have some multiplicative properties (hence their Z's have Euler-like products).

Lemma 4. *Let $c \in \mathbb{Z}_{>0}$, $k \in \mathbb{Z}$, $h|n$. Then*

$$S(k, nc, \epsilon^{-3}\rho_{n;h}) = \frac{-\mathrm{i}\sqrt{nc}}{2} \sum_{\substack{0\leq m<4nc \\ m^2\equiv_{8nc}1-8k+8c^2n/h}} e\left(\frac{m}{4nc}\right) . \tag{4.12}$$

Proof. We will begin with the proof of $n = h = 1$. Writing $m = 2\ell + 1$, define

$$B_c(k) := \frac{-\mathrm{i}\sqrt{c}}{2} \sum_{\substack{m=0 \\ m^2\equiv_{8c}1-8k}}^{4c-1} (-1)^{(m-1)/2} e\left(\frac{m}{4c}\right) = \frac{-\mathrm{i}\sqrt{c}}{2} e\left(\frac{1}{4c}\right) \sum_{\substack{\ell=0 \\ \ell^2+\ell\equiv_{2c}-2k}}^{2c-1} (-1)^{\ell} e\left(\frac{\ell}{2c}\right) .$$

As we manifestly have $B_c(k + c) = B_c(k)$ (i.e. B_c is a class function for $\mathbb{Z}_c$), we can formally write it as a combination of irreducible $\mathbb{Z}_c$-characters:

$$B_c(k) = \sum_{d=0}^{c-1} R_{d,c}\, e\left(\frac{dk}{c}\right) , \tag{4.13}$$

for coefficients

$$\begin{aligned} R_{d,c} &= \frac{-\mathrm{i}}{2\sqrt{c}} e\left(\frac{1}{4c}\right) \sum_{j=0}^{c-1} e\left(\frac{-dj}{c}\right) \sum_{\substack{\ell=0 \\ \ell^2+\ell\equiv_{2c}-2k}}^{2c-1} (-1)^{\ell} e\left(\frac{\ell}{2c}\right) \\ &= \frac{-\mathrm{i}}{2\sqrt{c}} e\left(\frac{1}{4c}\right) \sum_{\ell=0}^{2c-1} (-1)^{\ell} e\left(\frac{d\,(\ell^2+\ell)}{2c} + \frac{\ell}{2c}\right) . \end{aligned} \tag{4.14}$$

It is elementary to show that the sum on the right-side of (4.14) vanishes when $\gcd(c, d) > 1$: write $m = \gcd(c, d)$ and $\ell = s2c/m + r$, so $\sum_{\ell=0}^{2c-1} = \sum_{r=0}^{2c/m-1} \sum_{s=0}^{m-1}$, and notice that $\sum_s = 0$ for each r (when $m > 1$).

Thus we can restrict to $\gcd(c, d) = 1$. We want to show $R_{d,c} = \omega_{d,c}^{-3}$. Choose $d' \in \mathbb{Z}$ so that $1 = e\left(\frac{(dd'-1)(c+1)}{2c}\right)$; this permits us to rewrite (4.14) as

$$R_{d,c} = \frac{-\mathrm{i}e(1/4c)}{2\sqrt{c}} G(d, d\gamma; 2c) \tag{4.15}$$

where $\gamma := d'c + d' + 1$ and $G(a,b;c)$ is the generalised Gauss sum $G(a,b;c) = \sum_{\ell=0}^{c} e((a\ell^2 + b\ell)/c)$. Note that $\gamma := d'c + d' + 1$ is even iff c is even. When c is even, we can complete squares and obtain

$$\begin{aligned} R_{d,c} &= \frac{-\mathrm{i}\, e(1/4c)}{2\sqrt{c}}\, e\left(\frac{-d\gamma^2}{8c}\right) G(d,0;2c) \\ &= \left(\frac{2c}{d}\right) e\left(\frac{-(d-1)^2}{16} + \frac{2-c-d\gamma^2}{8c}\right). \end{aligned} \tag{4.16}$$

When c is odd, use $\frac{1}{2c} = \frac{-1}{2} + \frac{(c+1)/2}{c}$ to write $G(d, d\gamma; 2c) = G(-d, -d\gamma; 2)\, G(d(c+1)/2, d\gamma(c+1)/2; c)$; the left generalised Gauss sum equals 2, while the right is evaluated by completing squares as usual, and we obtain (for c odd)

$$R_{d,c} = \left(\frac{d(c+1)/2}{c}\right) e\left(\frac{(c-1)^2}{16} + \frac{2-2c-d(c+1)^3(d'+1)^2}{8c}\right). \tag{4.17}$$

Now, for m odd we have $\left(\frac{2}{m}\right) = (-1)^{(m^2-1)/8}$ and $\left(\frac{-1}{m}\right) = (-1)^{(m-1)/2}$. Consider first c even. Then directly from (4.3) and (4.16) we obtain

$$\frac{\omega_{d,c}^{-3}}{R_{d,c}} = e\left(\frac{-2-2c-d-d'-2cd+2dd'-c^2d+c^2d'+d^2d'+dd'^2+2cdd'^2+c^2dd'^2-c^2d^2d'}{8c}\right). \tag{4.18}$$

Because $dd' \equiv_{2c} 1$ when c is even we can define an integer L by $dd' = 1 + 2cL$, and we know that $d \equiv_4 d'$. Then (4.18) collapses to

$$\frac{\omega_{d,c}^{-3}}{R_{d,c}} = e\left(\frac{dL + d'L + 2L}{4}\right) = 1\,, \tag{4.19}$$

as desired. The proof for c odd is similar: the Jacobi symbol $\left(\frac{d(c+1)/2}{c}\right)$ equals $\left(\frac{2d}{c}\right)$; define L by $dd' = 1 + cL$ and use the congruences $c^3 \equiv_{8c} c$ and $4c^2 \equiv_{8c} 4c$.

Replacing c everywhere with nc establishes the $h = 1$ case of Lemma 4. Arbitrary $h|n$ is handled through the elementary observation that $S(k, nc, \epsilon^{-3}\rho_{n;h}) = S(k - c^2n/h, nc, \epsilon^{-3}\rho_{n;1})$. **QED to Lemma 4**

Hooley's method is based on the $n = 1$ case of Lemma 5 below; it was more recently modified slightly by [4]. We need a more serious revision. A binary quadratic form $Q(x,y) = \alpha x^2 + \beta xy + \gamma y^2 =: [\alpha, \beta, \gamma]$ is called *integral* if $\alpha, \beta, \gamma \in \mathbb{Z}$, and *positive-definite* if α and γ are both positive. (See e.g. Chapter 12 of [38] for a rather complete introduction to the basic theory of quadratic forms, as is relevant here.) The discriminant is $\beta^2 - 4\alpha\gamma$. For any $C \in \mathbb{Z}_{>0}$ and $D \in \mathbb{Z}_{<0}$, let $\mathcal{Q}(C, D)$ denote the set of all triples $(Q; r, s)$, where Q is integral and positive-definite with discriminant D and where r, s are coprime integers satisfying $Q(r,s) = C$. Any $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ acts on $(Q; r, s) \in \mathcal{Q}(C, D)$ by sending $Q(x,y)$ to $Q(ax + by, cx + dy)$ and (r, s) to $(dr - bs, -cr + as)$. In particular, it is easy to verify that if $(Q; r, s)$ and $(Q'; \tilde{r}, \tilde{s})$ are in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit then Q, Q' have the same discriminant and $Q(r,s) = Q'(\tilde{r},\tilde{s})$. Then Hooley observes that there is a bijection between the integers $0 \leq m < 2C$ satisfying $m^2 \equiv_{4C} 1$, and $\mathrm{SL}_2(\mathbb{Z})$-orbits in $\mathcal{Q}(C,D)$.

We need to generalise this in two ways. First, choose any integer $n \geq 1$. Write $\mathcal{Q}_n(C,D)$ denote the set of all triples $(Q;r,ns)$, where $Q = [n\alpha,\beta,\gamma]$ is positive-definite and of discriminant D, $\alpha,\beta,\gamma,r,s \in \mathbb{Z}$, $\gcd(r,ns) = 1$, and $Q(r,ns) = nC$. For example, $\mathcal{Q}(C,D) = \mathcal{Q}_n(C,D)$. It is elementary to show that the group $\Gamma_0(n)$ acts on $\mathcal{Q}_n(C,D)$.

Secondly, suppose that h divides $\gcd(n,24)$ and that $\gcd(n/h-1,h) = 1$. Write $n' = n/h$. Let $\mathcal{Q}_{n;h}(C,D)$ denote the set of all triples $(Q;r,ns)$ where $Q = [n\alpha,\beta,\gamma/h]$ is positive-definite and of discriminant D, $\alpha,\beta,\gamma,r,s \in \mathbb{Z}$, $\gamma \equiv_h \alpha$, $\gcd(r,n's) = 1$, and $Q(r,ns) = C$. Of course $\mathcal{Q}_{n;1}(C,D) = \mathcal{Q}_n(C,D)$. Recall the group $\Gamma_0(n|h)$, which can be defined equivalently as either the conjugate of $\Gamma_0(n')$ by $\left(\begin{smallmatrix} h & 0 \\ 0 & 1 \end{smallmatrix}\right)$, or as the set of all determinant-1 matrices of the form $\left(\begin{smallmatrix} a & b/h \\ nc & d \end{smallmatrix}\right)$ for $a,b,c,d \in \mathbb{Z}$. Let $\Gamma_0(n;h)$ denote the set of all $\left(\begin{smallmatrix} a & b/h \\ nc & d \end{smallmatrix}\right) \in \Gamma_0(n|h)$ for which $ac \equiv_h bd$. Then it is easy to verify $\Gamma_0(n;h)$ is a group, and we see shortly that it acts on $\mathcal{Q}_{n;h}(C,D)$.

The reason 24 arises here (and elsewhere) is because, for any divisor d of 24, any integer ℓ coprime to d satisfies $\ell^2 \equiv_d 1$. The defining condition $ac \equiv_h bd$ of $\Gamma_0(n;h)$ is equivalent to requiring that there is some ℓ (depending on a,b,c,d) coprime to h with $a \equiv_h \ell d$ and $b \equiv_h \ell$. To see this equivalence, run through each prime power p^ν exactly dividing h; the determinant condition $ad - n'bc = 1$ tells us that either a and d are both coprime to h (in which case require $\ell \equiv ad \pmod{p^\nu}$), or b,c are both coprime to h (in which case take $\ell \equiv bc \pmod{p^\nu}$). Then $a \equiv_{p^\nu} \ell d$ resp. $b \equiv_{p^\nu} \ell c$, and the other congruence comes from $ac \equiv_{p^\nu} bd$. Running through all p, we obtain an ℓ defined mod h which has all the desired properties.

Lemma 5(a) *Let $D \in \mathbb{Z}_{<0}$, $C \in \mathbb{Z}_{>0}$. There is a one-to-one correspondence between the set of integers m, $0 \leq m < 2nC$, satisfying $m^2 \equiv D \pmod{4nC}$, and $\Gamma_0(n)$-orbits in $\mathcal{Q}_n(C,D)$.*

(b) *There is a one-to-one correspondence between the set of integers m, $0 \leq m < 2nC$, satisfying $m^2 \equiv D + 4C^2n' \pmod{4nC}$, and $\Gamma_0(n;h)$-orbits in $\mathcal{Q}_{n;h}(C,D)$.*

Proof. Let's begin with the simpler part (a). First note that there is an obvious bijection between each $m \in \mathbb{Z}$ with $m^2 \equiv_{4nC} D$, and each quadratic form $Q = [nC,m,\gamma]$ (necessarily positive-definite) of discriminant D, where $\gamma = (m^2-D)/(4nC)$.

Given any $(Q;r,ns) \in \mathcal{Q}_n(C,D)$, $Q = [n\alpha,\beta,\gamma]$, choose any $\tilde{r},\tilde{s}$ such that $\begin{pmatrix} r & \tilde{r} \\ ns & \tilde{s} \end{pmatrix} \in \Gamma_0(n)$ (this is possible since $\gcd(r,ns) = 1$), and define the *root* $m := 2n\alpha r\tilde{r} + \beta\,(r\tilde{s} + ns\tilde{r}) + 2\gamma ns\tilde{s}$. Then $\begin{pmatrix} r & \tilde{r} \\ ns & \tilde{s} \end{pmatrix} \in \Gamma_0(n)$ sends $(Q;r,ns)$ to $([nC,m,\gamma'];1,0)$ where $\gamma' = (m^2-D)/(4nC)$. The other choices of $\tilde{r},\tilde{s}$ are $\tilde{r} + Lr, \tilde{s} + Lns$ for any $L \in \mathbb{Z}$, which correspond to roots $m + L2nC$, and so the root taken mod $2nC$ is a well-defined function $m(Q;r,ns) \in \mathbb{Z}_{2nC}$ on $\mathcal{Q}_n(C,D)$. The desired bijection in part (a) is this root map (mod $2nC$).

Now turn to part (b). Here, there is an elementary bijection between each $m \in \mathbb{Z}$ with $m^2 \equiv_{4nC} D + 4n'C^2$, and each positive-definite quadratic form $Q = [nC, m, \gamma/h]$ of discriminant D, where $\gamma \equiv_h C$. Choose any $([n\alpha, \beta, \gamma/h]; r, ns) \in \mathcal{Q}_{n;h}(C, D)$. Then $\left(\begin{smallmatrix} a & b/f \\ nc & d \end{smallmatrix}\right) \in \Gamma_0(n|h)$ will send (r, ns) to $(ra + n'sb, nrc + nsd)$ (the desired form), and $[n\alpha, \beta, \gamma/h]$ to

$$[n(\alpha a^2 + \beta ac + n'\gamma c^2), 2n'\alpha ab + \beta\,(ad + n'bc) + 2\gamma n'cd, (n'\alpha b^2 + \beta bd + \gamma d^2)/h]\,. \quad (4.20)$$

Then (4.20) will lie in $\mathcal{Q}_{n;h}(C, D)$ if

$$a^2 \equiv_h d^2\,, \quad ac \equiv_h bd\,, \quad n'c^2 \equiv_h n'b^2\,. \quad (4.21)$$

As explained above, the condition $ac \equiv_h bd$ is equivalent to the existence of an ℓ coprime to h satisfying $a \equiv_h \ell d$ and $b \equiv_h \ell c$, and such an ℓ forces the other two congruences to be satisfied. In other words, the matrices in $\Gamma_0(n|h)$ satisfying (4.21) form the group $\Gamma_0(n; h)$.

Since r and $n's$ are coprime, we can find integers $\tilde{r}, \tilde{s}$ such that $r\tilde{s} - n's\tilde{r} = 1$. We claim there is some $L \in \mathbb{Z}$, unique modulo h, such that $\left(\begin{smallmatrix} r & (Lr+\tilde{r})/h \\ ns & Ln's+\tilde{s} \end{smallmatrix}\right) \in \Gamma_0(n; h)$. To see this, choose any prime power p^ν exactly dividing h. If r is coprime to p choose $\ell \equiv_{p^\nu} 1 - n's^2$ and $L \equiv_{p^\nu} r'\,(-\tilde{r} + \ell s)$ where $r'r \equiv_{p^\nu} 1$; otherwise, $n's$ will be coprime to p, so choose s' by $s'n's \equiv_{p^\nu} 1$, $\ell_{p^\nu} r^2 - n'$, and $L \equiv_{p^\nu} s'\,(-\tilde{s} + \ell r)$. We are using here that $1 - n'$ is coprime to h and hence p. Then ℓ and L are defined mod h by running through all p. To see uniqueness of L modulo h, note that $\left(\begin{smallmatrix} r & (Lr+\tilde{r})/h \\ ns & Ln's+\tilde{s} \end{smallmatrix}\right)\left(\begin{smallmatrix} r & (L'r+\tilde{r})/h \\ ns & L'n's+\tilde{s} \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & (L-L')/h \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(n; h)$ forces $(L - L')/h \in \mathbb{Z}$ by definition of $\Gamma_0(n; h)$.

So we may assume $\left(\begin{smallmatrix} r & \tilde{r}/h \\ ns & \tilde{s} \end{smallmatrix}\right) \in \Gamma_0(n; h)$. Define the root $m := 2n'\alpha r\tilde{r} + \beta\,(r\tilde{s} + n's\tilde{r}) + 2\gamma n's\tilde{s}$ as before. Then $\begin{pmatrix} r & \tilde{r}/h \\ ns & \tilde{s} \end{pmatrix}$ sends $(Q; r, ns)$ to $([nC, m, \gamma']; 1, 0) \in \mathcal{Q}_{n;d}(C, D)$ where $\gamma' = (m^2 - D)/(4nC)$. From the uniqueness of the previous paragraph, we have that $\left(\begin{smallmatrix} r & (Lr+\tilde{r})/h \\ ns & Ln's+\tilde{s} \end{smallmatrix}\right) \in \Gamma_0(n; h)$ iff $L \in h\mathbb{Z}$, which corresponds to root $m + L2nC$, and so again the root taken mod $2nC$ is a well-defined function $m(Q; r, ns) \in \mathbb{Z}_{2nC}$ on $\mathcal{Q}_{n;d}(C, D)$. **QED to Lemma 5**

We are interested in $D = 1 - 8k$, $n = 2n_g$, $h = h_g$ and $C = c$. Restrict attention here to $k \geq 1$ (not a problem since we are only interested in large k), so $D < 0$. In fact, to sharpen slightly some of our bounds, we'll choose $k \geq 5$. Continue to write $n' = n/h$.

Note that if both $(Q; r, ns), (Q; \tilde{r}, n\tilde{s})$ (same Q) lie in $\mathcal{Q}_{n;k}(C, D)$, then there is an automorphism $g \in \Gamma_0(n; k)$ of Q sending (r, ns) to $(\tilde{r}, n\tilde{s})$. For $D < 0$ (the case we are interested in), each such automorphism g must have finite order (since completing squares in $\alpha = \alpha a^2 + \beta ac + n'\gamma c^2$ and $\gamma = n'\alpha b^2 + \beta bd + \gamma d^2$ bounds the matrix entries a, b, c, d of g). Now, $\Gamma_0(n|h)$ (which contains $\Gamma_0(n; h)$) is conjugate to

$\Gamma_0(n')$, so all of its elements of finite order have orders 2, 4, 6. In fact for us, the stabiliser $\mathrm{stab}([n\alpha,\beta,\gamma/h];\Gamma_0(n;h))$ of $[n\alpha,\beta,\gamma/h]$ in $\Gamma_0(n;h)$ will always be ± 1. To see this, first note $\mathrm{stab}([n\alpha,\beta,\gamma/h];\Gamma_0(n;h))$ is a subgroup of $\mathrm{stab}([n\alpha,\beta,\gamma/h];\Gamma_0(n|h))$, which is isomorphic to $\mathrm{stab}([n'\alpha,\beta,\gamma];\Gamma_0(n;h))$, which is in turn a subgroup of $\mathrm{stab}([n'\alpha,\beta,\gamma];\mathrm{SL}_2(\mathbb{Z}))$. For a *primitive* form $[n'\alpha,\beta,\gamma]$ (i.e. when $n'\alpha,\beta,\gamma$ don't have a common factor), nontrivial stabilisers occur only for discriminant $D=-3,-4$. But our discriminant satisfies $D\equiv 1 \pmod 8$, so even if $[n'\alpha,\beta,\gamma]$ is imprimitive, it can never have a nontrivial stabiliser in $\mathrm{SL}_2(\mathbb{Z})$.

We also need a bound on $\Gamma_0(n;h)$-equivalence class representatives $[n\alpha,\beta,\gamma/h]$. For this purpose, observe that $\Gamma_0(n;h)$ contains $\Gamma_0(nh)$, which has finite index $nh\prod_p(1+1/p)$ in $\mathrm{SL}_2(\mathbb{Z})$. The cosets for $\Gamma_0(N)\backslash\mathrm{SL}_2(\mathbb{Z})$ are in bijection with pairs $(c,d)\in\mathbb{Z}^2_{>0}$ where $c|N$ and $1\leq d\leq N/c$ satisfies $\gcd(d,c,N/c)=1$. To any such pair (c,d), a representative of that coset is $\left(\begin{smallmatrix} a & b \\ c & d'\end{smallmatrix}\right)$ where d' is coprime to c and satisfies $d'\equiv_{N/c} d$, and a,b are any integers satisfying $ad'-bc=1$. Now, $c=N$ corresponds to the identity coset, so for it take $c=0$ instead; then in all cases we have $c\leq N/2$. We can always choose $|a|\leq c/2$ by adjusting b appropriately (at least when $c\neq 0$), hence $|a|\leq N/4$ (true even for $c=0$ unless $N<4$). Now, over $\mathrm{SL}_2(\mathbb{Z})$, any positive-definite discriminant D quadratic form (with coefficients in $h^{-1}\mathbb{Z}$) is equivalent to some $[\alpha',\beta',\gamma']$ (namely Gauss' reduced form) with $|\beta'|\leq\alpha'\leq\gamma'$, where $\alpha'\leq\sqrt{|D|/3}$ and $\gamma'\leq 11h|D|/39$. Combining with (4.20), and noting that we can always force $|\beta|\leq n\alpha$ by applying multiples of $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1\end{smallmatrix}\right)\in\Gamma_0(nh)$, we see that any such quadratic form is equivalent over $\Gamma_0(nh)$, hence over $\Gamma_0(n;h)$, to some $[n\alpha,\beta,\gamma/h]$ satisfying the (crude but adequate) bounds

$$0<\alpha\leq\frac{n^2h^2}{16n}\left(3\sqrt{\frac{|D|}{3}}+44h|D|/39\right)<\frac{1}{12}nh^3\,|D|\,, \tag{4.22}$$

$$|\beta|<n^2h^3\,|D|/12\,, \tag{4.23}$$

$$0<\gamma\leq(|D|+\beta^2)/(4n'\alpha)<\begin{cases}|D|/7 & \text{if } n=2\\ n^2h^4\,|D|/44 & \text{otherwise}\end{cases}. \tag{4.24}$$

(The given bound on α, hence β, is also true for the identity coset $c=0$.) Implicit in these derivations are $2h\leq n$, n even, and $|D|\geq 39$.

The point is that the series $Z_{n;h}(3/4)$ can be rewritten as a sum over the finitely many $\Gamma_0(n;h)$-orbit representatives $[n\alpha,\beta,\gamma/h]$, and over integers r,s. Over-counting by a factor of 2, we can take this to be a free sum over r,s.

Theorem 3. *Choose any $k,n\in\mathbb{Z}_{>0}$, $k\geq 5$, and any h dividing gcd$(n,24)$ such that gcd$(n/h-1,2h)=1$. Then for $n\neq 2$,*

$$\begin{aligned}\left|Z_{n/2;h}(3/4)\right|\leq{}&\left(\prod_{p|(nh)}\frac{p+1}{p}\right)\left(1+2.13|D|^{1/8}\log|D|\right)\times\\ &\big((6.124n^{35/6}h^{47/6}-3.09n^{23/4}h^{31/4}+64.32n^{29/6}h^7-23n^{19/4}h^7)|D|\\ &+(.146n^{47/6}h^{65/6}-.114n^{31/4}h^{43/4}+2.51n^{35/6}h^{10}-.74n^{23/4}h^{10})|D|^{3/2}\big)\ .\end{aligned}$$

where we write $D = 1-8k$*. For* $n = 2$*, this bound on the right-side should be replaced with* $(3872|D| + 213|D|^{3/2})(1 + 2.13|D|^{1/8}\log|D|)$*.*

Proof. Note the n in the statement of Theorem 3 agrees with that of Lemma 5, but is twice that of Lemma 4. We need to bound the limit as $X \to \infty$ of

$$Z_{n/2;h}(s;X) = \sum_{c=1}^{X} S(k, nc, \epsilon^{-3}\rho_{n/2;h})\,(nc/2)^{-2s}\,. \tag{4.25}$$

Thanks to Lemmas 4 and 5, we can write its value at $s = 3/4$ as

$$\frac{-\mathrm{i}}{2}\sum_{[n\alpha,\beta,\gamma/h]}\ \sum_{c(\alpha,\beta,\gamma;r,s)\leq X}{}' \frac{(-1)^{(m(\alpha,\beta,\gamma;r,s)-1)/2}}{nc(\alpha,\beta,\gamma;r,s)}\, e\left(\frac{m(\alpha,\beta,\gamma;r,s)}{2n\,c(\alpha,\beta,\gamma;r,s)}\right)\,, \tag{4.26}$$

where the first sum is over representatives of the $\Gamma_0(n;h)$-orbits of positive-definite quadratic forms of discriminant $D = 1-8k$, and the second sum is over all integers r, s satisfying $\gcd(r, n's) = 1$ and the given inequality (the prime on the sum denotes that coprime condition). The quantities m, c are defined by

$$m(\alpha,\beta,\gamma;r,s) = 2n'\alpha r\tilde{r} + \beta\,(r\tilde{s} + n's\tilde{r}) + 2\gamma n's\tilde{s}\,,\ \ c(\alpha,\beta,\gamma;r,s) = \alpha r^2 + \beta rs + n'\gamma s^2\,,$$

where $\tilde{r}, \tilde{s}$ are any integers satisfying $\left(\begin{smallmatrix} r & \tilde{r}/h \\ ns & \tilde{s}\end{smallmatrix}\right) \in \Gamma_0(n;h)$ — which pair is chosen won't affect the value of that summand. The factor of 2 in (4.26) comes from the aforementioned redundancy that different (r, ns) can lie in the same orbit.

Write this inner sum $\sum_{r,s}$ (for each choice of α, β, γ) as $\sum_{>} + \sum_{<}$, depending on whether or not $|r| > n|s|$ ($|r| = n|s|$ would contradict $\gcd(r,s) = 1$) . We will bound $Z_{n/2;h}(3/4;X)$ by considering separately the contributions of $\sum_{>}, \sum_{<}$.

The arguments for $\sum_{>}$ and $\sum_{<}$ are completely analogous, and so we will consider in detail only $\sum_{>}$. Since then $r > 0$, we can write

$$\frac{m}{2nc} = \frac{2n'\tilde{r}\,(\alpha r^2 + \beta rs + n'\gamma s^2) + 2n'\gamma s + \beta r}{2nr\,(\alpha r^2 + \beta rs + n'\gamma s^2)} \equiv_1 -\frac{us'}{R} + \frac{2n'\gamma s + \beta r}{2nr\,(\alpha r^2 + \beta rs + n'\gamma s^2)} \tag{4.27}$$

where $\delta = \gcd(s,h)$, $h' = h/\delta$, $R = rh'$, $s' \in \mathbb{Z}$ is any inverse of s/δ mod R, and $u \in \mathbb{Z}$ is defined mod R by $u \equiv_r -\delta^{-1}(n')^{-1}$, $u \equiv_{h'} 1 - s^2n'$, and (if $3|\gcd(r,h)$) $u \equiv_9 -(n')^{-1}$ (note that $\gcd(r,h) \neq 1$ implies it equals 3, in which case s is coprime to 3 and $n' \equiv_3 -1$). Writing

$$\alpha r^2 + \beta rs + n'\gamma s^2 = n'\gamma\,(s + \frac{\beta r}{2n'\gamma})^2 - \frac{Dr^2}{4n'\gamma} = \alpha\,(r + \frac{\beta s}{2\alpha})^2 - \frac{Ds^2}{4\alpha}\,, \tag{4.28}$$

the first inequality gives

$$|r| \leq \sqrt{\frac{X4n'\gamma}{|D|}} = C\sqrt{X}\,, \qquad S_- \leq s \leq S_+\,, \tag{4.29}$$

where $C = \sqrt{4n'\gamma/|D|}$ and

$$S_+ = \min\left\{\frac{|r|}{n}, -\frac{\beta r}{2n'\gamma} + \frac{1}{2n'\gamma}\sqrt{4n'\gamma X + Dr^2}\right\}, \tag{4.30}$$

$$S_- = \max\left\{-\frac{|r|}{n}, -\frac{\beta r}{2n'\gamma} - \frac{1}{2n'\gamma}\sqrt{4n'\gamma X + Dr^2}\right\}. \tag{4.31}$$

We compute

$$(-1)^m = (-1)^{n'(\alpha r\tilde{r} + \beta s\tilde{r} + \gamma s\tilde{s}) + (\beta-1)/2} = (-1)^{(\beta-1)/2} \tag{4.32}$$

using the evenness of n' and the determinant relation $r\tilde{s} - n's\tilde{r} = 1$ (β is odd because of the discriminant condition $\beta^2 - 4n'\alpha\gamma = 1 - 8k$). Now put

$$\varphi(r,s) = \frac{(-1)^{(\beta-1)/2}}{\alpha r^2 + \beta rs + \gamma n's^2}\, e\left(\frac{2n'\gamma s + \beta r}{2nr\left(\alpha r^2 + \beta rs + \gamma n's^2\right)}\right); \tag{4.33}$$

we need to bound

$$\sum_{>} = \sum_{|r| \le C\sqrt{X}} \sum_{\delta | h} \sideset{}{'}\sum_{\substack{S_- \le s \le S_+ \\ \gcd(s,h)=\delta}} e\left(\frac{us'}{hr/\delta}\right)\varphi(r,s) = \sum_{|r| \le C\sqrt{X}} \sum_{\delta | h} \sum_{\substack{S_-/\delta \le S \le S_+/\delta \\ \gcd(S,R)=1}} e\left(\frac{uS'}{R}\right)\varphi(r,\delta S), \tag{4.34}$$

where we use (4.27), we write $R = hr/\delta$ as before and $S' = s'$ is any inverse of S mod R. Rewrite the sum over S in (4.34) (for fixed $\alpha, \beta, \gamma, r, \delta$) using partial summation (the discrete analogue of integration by parts):

$$\begin{aligned}\sideset{}{'}\sum_{S_-/\delta \le S \le S_+/\delta} e\left(\frac{uS'}{R}\right)\varphi(r,\delta S) &= \sum_{S_-/\delta \le \sigma \le S_+/\delta} g(\sigma)\left(\varphi(r,\delta\sigma) - \varphi(r,\delta\sigma+\delta)\right) \\ &\quad + g(\lfloor S_+/\delta \rfloor)\,\varphi(r,\delta\lfloor S_+/\delta\rfloor),\end{aligned} \tag{4.35}$$

where

$$g(\sigma) = \sideset{}{'}\sum_{\substack{S_-/\delta \le S \le \sigma \\ \gcd(S,R)=1}} e(uS'/R) \tag{4.36}$$

denotes the *incomplete* Kloosterman sums (R is implicit). We have (for $n > 2$)

$$\begin{aligned}|\varphi(r,s) - \varphi(r,s+\delta)| &\le \frac{16n'^2\gamma^2\delta}{D^2r^4}\left|\frac{|r|}{2}\sqrt{|D|} + n'\gamma\delta\right| + \frac{2\pi}{2n|r|}\frac{(4n'\gamma)^2}{|D|r^2}\frac{2}{2|r|\sqrt{|D|}} \\ &\le .0042\frac{n^6h^6}{|r|^3}\delta\sqrt{|D|} + .000188\frac{n^9h^9}{r^4}\delta^2|D| + .0083\frac{n^5h^6}{r^4}\sqrt{|D|},\end{aligned} \tag{4.37}$$

$$|\varphi(r,\delta\lfloor S_+/\delta\rfloor)| \le 4n'\gamma/(|D|r^2) \le .091n^3h^3r^{-2}, \tag{4.38}$$

where we use repeatedly (4.28), as well as Taylor's remainder $|e(\theta) - 1| \le 2\pi\theta_0$ for $|\theta| \le \theta_0$, and the elementary bound $x/(x^2+a^2) \le 1/(2|a|2)$. Thus all that remains is to bound $g(\sigma)$. The order of its growth with $|R|$ is stated (without proof) in Lemma 3 of [37]; an effective bound is:

Lemma 6. *For any integers k, u, h_1, h_2 with $h_1 \leq h_2$, we have*

$$\left|\sum_{h_1\leq h\leq h_2}{}' e\left(\frac{uh'}{k}\right)\right| < \left(\frac{k+h_2-h_1}{k}+2+2\ln(k)\right)\sqrt{k}\sqrt{\gcd(u,k)}\,d(k)\,, \tag{4.39}$$

where h' denotes the inverse of h (mod k) and the prime over the summation means to restrict to gcd$(h,k)=1$.

Proof. Assume for now that $h_2-h_1<k$. For any integer h, $h_1\leq h\leq h_1+k-1$, the map sending h to

$$\left(\left(\frac{h-h_2-1/2}{k}\right)\right)-\left(\left(\frac{h-h_1+1/2}{k}\right)\right)+\frac{h_2-h_1+1}{k}=\left\lfloor\frac{h-h_1+1/2}{k}\right\rfloor-\left\lfloor\frac{h-h_2-1/2}{k}\right\rfloor$$

equals 1 for $h_1\leq h\leq h_2$ and 0 otherwise. Now, we claim that for $0<x<1$,

$$\left|((x))+\sum_{j=1}^{k}\frac{\sin(2\pi jx)}{\pi j}\right|<\frac{1}{2k\min(x,1-x)}\,. \tag{4.40}$$

To see this, it suffices to consider $0<x\leq 1/2$. Write $K(x)=1+2\sum_{j=1}^{k}\cos(2\pi jx)=\sin(\pi\,(2k+1)x)/\sin(\pi x)$; then

$$((x))+\sum_{j=1}^{k}\frac{\sin(2\pi jx)}{\pi j}=\int_0^x K(y)dy-1/2=-\int_x^{1/2}K(y)dy \tag{4.41}$$

$$=\frac{-1}{\pi\,(2k+1)}\frac{\cos(\pi\,(2k+1)x)}{\sin(\pi x)}+\frac{1}{2k+1}\int_x^{1/2}\frac{\cos(\pi\,(2k+1)y)}{\sin^2(\pi y)}dy\,. \tag{4.42}$$

Since $\sin(\pi t)\geq 2t$ for $0\leq t\leq 1/2$, we get

$$\left|((x))+\sum_{j=1}^{k}\frac{\sin(2\pi jx)}{\pi j}\right|\leq\frac{1}{2\pi\,(2k+1)x}+\frac{1}{4\,(2k+1)x} \tag{4.43}$$

which implies the weaker (4.40).

Write $S(v,u;k)$ for the (complete) Kloosterman sum $\sum'_{1\leq h\leq k}e((vh+uh')/k)$ where the prime denotes restricting the sum to $\gcd(h,k)=1$. Then e.g. Lemma 2 of [36] gives an effective bound for it:

$$|S(v,u;k)|\leq\sqrt{k}\sqrt{\gcd(u,k)}\,d(k)\,. \tag{4.44}$$

Finally, it is elementary that $\sum_{j=1}^{k} 1/j \leq 1 + \ln(k)$. Therefore, putting all this together, we obtain

$$\left|{\sum_{h=h_1}^{h_2}}' e\left(\frac{h'}{k}\right)\right| \leq \frac{h_2 - h_1 + 1}{k}|S(0,u;k)| + \left|\sum_{j=1}^{k} \frac{e(-(k+1/2)j/k) - e(-j/(2k))}{2\pi j} S(j,u;k)\right|$$
$$+ 2\sum_{j=1}^{k/2} \frac{1}{2k\,(j-1/2)/k} \leq \frac{h_2 - h_1 + 1}{k}\sqrt{k}\sqrt{\gcd(u,k)}\, d(k)$$
$$+ \frac{1}{\pi}\sqrt{k}\sqrt{\gcd(u,k)}\, d(k)(1+\ln(k)) + (1+\ln(k))\,. \tag{4.45}$$

That the inequality (4.39) also holds when $h_2 - h_1 \geq k$ now follows from (4.44).
QED to Lemma 6

Effective bounds for the divisor function $d(n) = \sigma_0(n)$ are $d(n) \leq C_\epsilon n^\epsilon$, for any $\epsilon > 0$, where

$$C_\epsilon = \prod_{p < 2^{1/\epsilon}} \frac{1}{\epsilon\, \ln(p)\, e^{1-\epsilon \ln(p)}}\,, \tag{4.46}$$

where the product runs over all primes $p < 2^{1/\epsilon}$. To see this, recall $d(n) = \prod_p (a_p + 1)$ when $n = \prod_p p^{a_p}$ is the prime decomposition, so $d(n)/n^\epsilon = \prod_p \frac{a_p+1}{p^{\epsilon a_p}}$. The primes appearing in (4.46) are precisely those for which $\frac{a_p+1}{p^{\epsilon a_p}}$ can be > 1; the power a_p is then chosen to maximise this factor. Any $\epsilon < 1/2$ works for us, e.g. $C_{1/4} \approx 9.11...$ (in fact a slightly more refined analysis shows $C_{1/4}$ can be taken to be 8.55). Also, from $\ln y \leq y - 1$ we obtain $\ln(k) < 12\, k^{1/12} - 12$. Therefore we obtain as a bound on (4.35):

$$\left|\sum_{S_-/\delta \leq S \leq S_+/\delta} e\left(\frac{uS'}{R}\right)\varphi(r,\delta S)\right|$$
$$\leq (103\left|\frac{rh}{\delta}\right|^{5/6} - 74.8\left|\frac{rh}{\delta}\right|^{3/4})(.0084\frac{n^5h^6}{r^2}\sqrt{|D|} + .000377\frac{n^8h^9}{|r|^3}\delta|D| + .0164\frac{n^4h^6}{\delta|r|^3}\sqrt{|D|} + \frac{.091n^3h^3}{r^2})$$
$$< (1.065\frac{n^5h^{41/6}}{\delta^{5/6}|r|^{7/6}} - .733\frac{n^5h^{27/4}}{\delta^{3/4}|r|^{5/4}})\sqrt{|D|} + (.0389\frac{n^8h^{59/6}\delta^{1/6}}{|r|^{13/6}} - .028\frac{n^8h^{39/4}\delta^{1/4}}{|r|^{9/4}})|D|\,. \tag{4.47}$$

We can now bound the limit of (4.34) as $X \to \infty$, using the inequalities $n^{1-\mu}/(\mu-1) < \sum_{r=n}^{\infty} r^{-\mu} < n^{-\mu} + n^{1-\mu}/(\mu-1)$ for $\mu > 1$ which follow by comparing series with integral:

$$\lim_{X\to\infty}\left|\sum_{>}\right| \leq (19.136n^{29/6}h^{41/6} - 9.67n^{19/4}h^{27/4})\sqrt{|D|} + (.455n^{41/6}h^{59/6} - .359n^{27/4}h^{39/4})|D|\,. \tag{4.48}$$

When $n = 2$ (so $h = \delta = 1$), we get instead the weaker $\lim_{X\to\infty} |\sum_{>}| < 2136\sqrt{|D|} + 61|D|$.

Bounding $\sum_{<}$ is completely analogous; (4.27) becomes

$$\frac{m}{2nc} \equiv_1 \frac{ur'}{S} - \frac{\alpha r + \beta s}{2ns\,(\alpha r^2 + \beta rs + \gamma n' s^2)}\,, \tag{4.49}$$

where $\delta = \gcd(r, h) \in \{1, 3\}$, $S = ns/\delta$, r' is the inverse of r/δ mod S, and u is defined mod S by $u \equiv_S \delta^{-1}\,(1 + n's^2 - n'^2 s^4)$. The other equations and bounds are obtained by interchanging $r \leftrightarrow s$ and $\alpha \leftrightarrow n'\gamma$. In (4.34) and elsewhere it suffices to take $\delta \in \{1, 3\}$. Then (4.37),(4.38) become

$$|\varphi(r, s) - \varphi(r + \delta, s)| \le .056\frac{n^2h^6\delta}{|s|^3}\sqrt{|D|} + .0093\frac{n^3h^9\delta^2}{s^4}|D| + .34\frac{nh^6}{s^4}\sqrt{|D|}\,, \tag{4.50}$$

$$|\varphi(\delta\lfloor S_+/\delta\rfloor, s)| \le nh^3s^{-2}/3\,. \tag{4.51}$$

Therefore (4.35) becomes

$$\left|\sum_{S_-/\delta\le R\le S_+/\delta} e\left(\frac{uR'}{S}\right)\varphi(\delta R, s)\right| < (1.916\frac{n^{29/6}\delta^{1/6}}{|s|^{13/6}} - 1.26\frac{n^{19/4}\delta^{1/4}}{|s|^{9/4}})h^9|D|$$
$$+ (12.98\frac{n^{23/6}h^6}{\delta^{5/6}|s|^{7/6}} + 70.1\frac{n^{17/6}h^6}{\delta^{11/6}|s|^{13/6}} - 8.56\frac{n^{15/4}h^6}{\delta^{3/4}|s|^{5/4}} - 46\frac{n^{11/4}h^6}{\delta^{7/4}|s|^{9/4}})\sqrt{|D|}\,,$$

and we obtain

$$\lim_{X\to\infty}\left|\sum_{<}\right| \le (7.832n^{29/6}h^9 - 2.33n^{19/4}h^9)|D| + (201n^{23/6}h^6 - 72n^{15/4}h^6)\sqrt{|D|}\,.$$

All that remains is the sum over the representatives $[n\alpha, \beta, \gamma/h]$ of $\Gamma_0(n; h)$-equivalence classes of quadratic forms of discriminant D. Let $h'_{n;h}(D)$ denote the number of those (not necessarily primitive) equivalence classes. Note that our sum is now independent of α, β, γ, so a bound for $|Z_{n/2;h}(3/4)|$ will be the sum of our bounds for $\sum_{>}$ and $\sum_{<}$, multiplied by $h'_{n;h}(D)/2$ (the 2 here compensates for the overcounting, as mentioned earlier). Now, since all stabilisers here are $\pm I$, $h'_{n;h}(D)$ will equal the number of $\Gamma_0(n')$-equivalence classes of forms $[n'\alpha, \beta, \gamma]$, times the index of $\Gamma_0(n; h)$ in $\Gamma_0(n|h)$, and so we obtain the (crude) bound

$$h'_{n;h}(D) \le h'(D)\left|\frac{\mathrm{SL}_2(\mathbb{Z})}{\Gamma(n, h)}\right| \tag{4.52}$$

where $h'(D)$ is the total number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of integral forms (not necessarily primitive) of discriminant D, and where $\Gamma(n, h)$ consists of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ where n divides c and h divides b (this clearly forms a group). In deriving (4.52), we are conjugating everything by $\left(\begin{smallmatrix} h & 0 \\ 0 & 1 \end{smallmatrix}\right)$; $\Gamma(n, h)$ arises as the conjugate of the subgroup $\Gamma_0(nh)$ of $\Gamma_0(n;h)$. Now, the index of $\Gamma(n,h)$ in $\mathrm{SL}_2(\mathbb{Z})$ equals that of its conjugate $\Gamma_0(nh)$, since indices of Fuchsian groups can be expressed as a ratio of areas of fundamental domains and conjugating by $\mathrm{SL}_2(\mathbb{R})$ preserves those areas. Let $h(D)$ denote the class number of D, i.e. the number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of *primitive* integral quadratic forms of discriminant D. Then we have the bound

$$h(D) < \frac{\sqrt{|D|}}{\pi}\,(2+\log|D|) \tag{4.53}$$

(see Theorems 12.14.3 and 12.10.1 in [38]). Hence

$$\begin{aligned}\frac{h'(D)}{2} = \frac{1}{2}\sum_{m^2|D} h(D/m^2) &< \frac{\sqrt{|D|}}{2\pi}\sum_{m^2|D}\frac{2-\log m}{m} + \sqrt{|D|}\frac{\log|D|}{4\pi}\sum_{m^2|D} 1\\ &< 0.32\sqrt{|D|} + .681|D|^{5/8}\log|D|\,, \end{aligned}\tag{4.54}$$

using our earlier bound for the divisor function $d(n)$. Putting this all together, we obtain

$$h'_{n;h}(D)/2 < (0.32\sqrt{|D|} + .681|D|^{5/8}\log|D|)nh\prod_{p|(nh)}\frac{p+1}{p}\,, \tag{4.55}$$

and hence the bound given in the statement of our theorem. **QED to Theorem 3**

We can approximate the $2.13|D|^{1/8}\log|D|$ with the upper bound $120(|D|^{1/7}-|D|^{1/8})$ by the usual reasoning. Putting this all together, we obtain the following (very crude) bounds for the character values for each $g\in M_{24}$ and $k\geq 150$:

$$H_k(1) > \frac{4}{K}e^{\pi K/2} - \frac{4\pi}{\sqrt{2}}e^{\pi K/4} - 2.5\times 10^4 K^{23/7}\,, \tag{4.56}$$

$$|H_k(g)| < \frac{4}{K\sqrt{n_g}}e^{\pi K/(2n_g)} + \sqrt{8}\pi\, e^{\pi K/(4n_g)} + a_g\times 10^{b_g}K^{23/7}\,, \tag{4.57}$$

where $n_g=|g|$ is the order of g, $K=\sqrt{8k-1}$, and the values a_g, b_g are computed from Theorem 3 and are collected in Table 2. We find that when $k\geq 390$, the ratio $(H_k(1)/|M_{24}|)/(|H_k(g)|/|C_{M24}(g)|)$ exceeds 1.6×10^5 for all elements $g\neq 1$ except for g in conjugacy class 12B, where the ratio is >1.3. In particular we find that

$$\frac{H_k(1)}{|M_{24}|} > \sum_{g\neq 1}\frac{|H_k(g)|}{|C_{M24}(g)|}\,, \tag{4.58}$$

for all $k\geq 390$, where the sum is over all conjugacy classes in M_{24}. As explained in (4.1), this is sufficient to deduce positivity of the multiplicities of all M_{24}-irreps in these class functions H_k. Again, positivity (or rather nonnegativity) has been obtained experimentally for n up to 500, far more than we need.

It is somewhat disappointing that we need to check by hand (or rather, by computer!) positivity for so many k, when empirically (4.58) is satisfied for $k>30$. The cause of this are the bounds of Theorem 3, which apparently are far looser than they could be. But the value gained in tightening these bounds is far less than the effort to be spent, since it is so easy to calculate these character multiplicities. The reader interested in reducing 390 should focus on improving the bound for the case 12B (i.e. fix $n = 24, h = 12$ from the start, and restrict to $k > 100$ say; the index of $\Gamma_0(nh)$ in $\Gamma_0(n; h)$ here is 12, so the Theorem 3 bound can be improved immediately by that factor 12).

The bounds (4.56),(4.57) imply:

Theorem 4. *Choose any M_{24}-irrep ρ. Then the multiplicity $mult_\rho(H_n)$ of ρ in H_n tends to ∞ as $n \to \infty$. Moreover, $mult_\rho(H_n) > 0$ for all $n \geq 25$.*

Indeed, in this section we proved positivity of the multiplicities for $n \geq 390$; the values $25 \leq n < 390$ can be checked explicitly. An immediate corollary of Theorem 4 is the validity of Conjecture 5.11 in Umbral Moonshine [7]. However it should be remarked that the proof of this in [12] actually established a much stronger statement. Oddness of McKay-Thompson coefficients is far less trivial than strict positivity of certain multiplicities, since as we see almost every multiplicity will be strictly positive.

5 Is the Conway group the stringy symmetry?

In beautiful work, [30] followed Kondo's treatment [39] of Mukai's classification [42] of symplectic automorphisms of K3 surfaces, to obtain the symmetries of K3 sigma models. (Similar considerations are considered in [50].) It turns out all these stringy symmetries are subgroups of the Conway groups Co_0 or Co_1 (but not M_{24}), where they will necessarily generate the full group Co_0 (resp. Co_1). This begs the (perhaps naive) question: should the automorphism group of Mathieu Moonshine actually be a Conway group?

The Conway group Co_0 is the automorphism group of the Leech lattice; the simple group Co_1 is its quotient by its centre ± 1. Hence the Co_1-irreps are a subset of those of Co_0, consisting of those Co_0-irreps whose kernel contains -1. We call a Co_0-irrep a *spinor* if it is not trivial on -1. M_{24} is a subgroup of both Co_0 and Co_1. In this section we give the restrictions of Co_0 (hence Co_1) irreps to M_{24}, for irreps with dimension up to 1 million (there are 32 of these). The notation comes from the Atlas of Finite Groups [8]; in a couple places it differs slightly from that of [24] (e.g. the dimension-1035 irreps are in a different order: the conjugate ones we write as $1035'$ and $1035''$). The first step was identifying the conjugacy classes; once these are known, the character multiplicities follow directly from the chararacter table of M_{24} together with the orthogonality relations. The arguments are straightforward and we'll avoid the details, giving only the results.

Table 5. Matching of conjugacy classes

M_{24}	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	23A	23B
Co_1	1A	2A	2C	3B	3D	4D	4C	4F	5B	6E	6I	7B	7B	8E	10F	11A	12J	12M	14A	14B	15D	15D	21C	21C	23A	23B

Table 6. The 32 smallest Conway irreps

$\widetilde{\sigma}_1$	σ_1	σ_2	σ_3	$\widetilde{\sigma}_2$	$\widetilde{\sigma}_3$	$\widetilde{\sigma}_4$	σ_4	σ_5	σ_6	σ_7	$\widetilde{\sigma}_5$	σ_8	σ_9	σ_{10}	$\widetilde{\sigma}_6$	$\widetilde{\sigma}_7$	$\widetilde{\sigma}_8$	σ_{11}	$\widetilde{\sigma}_9$	σ_{12}	$\widetilde{\sigma}_{10}$	σ_{13}	$\widetilde{\sigma}_{11}$	σ_{14}	σ_{15}	σ_{16}	σ_{17}	$\widetilde{\sigma}_{12}$	σ_{18}	σ_{19}
24	276	299	1771	2024	2576	4576	8855	17250	27300	37674	40480	44275	80730	94875	95680	170016	299000	313950	315744	345345	351624	376740	388080	483000	644644	673750	673750	789360	822250	871884

We find that any *virtual character* $\rho = m_0 + m_1\rho_1 + \cdots + m_{20}\rho_{20}$ of M_{24}, i.e. any linear combination with integer coefficients m_i of M_{24} irreps, is the restriction of a virtual Co_0-representation involving the first 32 irreps of Co_0, iff the multiplicities obey $m_2 = m_{\overline{2}}$, $m_3 = m_{\overline{3}}$, $m_8 = m_{\overline{8}}$, $m_{10} = m_{\overline{10}}$, as well as the relation $m_9 + m_{14} \equiv_2 m_{15}$. Thus it is hard for an M_{24}-representation to be a restriction of a virtual Co_0-representation (0% will be). Although the restriction map is not surjective, Theorem B tells us the Mathieu-Moonshine characters H_n lie in the image of that restriction. This is the content of Theorem C given earlier. In particular we have:

$$\begin{aligned}
&\rho_1 = \widetilde{\sigma}_1 - 1\,,\\
&\rho_2 + \overline{\rho_2} = 37\sigma_{14} + 39\sigma_{13} + 28\widetilde{\sigma}_{10} + \sigma_{12} + 12\sigma_{11} + 23\widetilde{\sigma}_8 + \sigma_{10} + 25\sigma_9 + 2\sigma_5 + 194\widetilde{\sigma}_4 + 24\widetilde{\sigma}_3 + 18\sigma_3\\
&\qquad +269\widetilde{\sigma}_1 - 28\sigma_{18} - 11\widetilde{\sigma}_{12} - 42\widetilde{\sigma}_9 - 44\widetilde{\sigma}_6 - 9\sigma_8 - 100\widetilde{\sigma}_5 - 18\sigma_7 - 78\sigma_6 - 144\sigma_2 - 218\sigma_1 - 155\,,\\
&\rho_3 + \overline{\rho_3} = 14\sigma_{13} + 11\widetilde{\sigma}_{10} + 15\sigma_{12} + 4\sigma_{11} + 6\widetilde{\sigma}_8 + 15\sigma_{10} + 8\sigma_9 + 105\widetilde{\sigma}_4 + 14\widetilde{\sigma}_3 + 64\sigma_3 + 121\widetilde{\sigma}_1\\
&\qquad -10\sigma_{18} - 4\widetilde{\sigma}_{12} - 15\widetilde{\sigma}_9 - 15\widetilde{\sigma}_6 - 20\sigma_8 - 12\widetilde{\sigma}_5 - 20\sigma_7 - 13\sigma_6 - 97\sigma_2 - 49\sigma_1 - 27\,,\\
&\rho_4 = \sigma_2 + 1 - 2\widetilde{\sigma}_1\,,\\
&\rho_5 = \sigma_1 + 1 - \widetilde{\sigma}_1\,,\\
&\rho_6 = 8\sigma_{14} + 14\sigma_{13} + 11\widetilde{\sigma}_{10} + 7\sigma_{12} + 4\sigma_{11} + 6\widetilde{\sigma}_8 + 7\sigma_{10} + 8\sigma_9 + 89\widetilde{\sigma}_4 + 14\widetilde{\sigma}_3 + 32\sigma_3 + 121\widetilde{\sigma}_1\\
&\qquad -10\sigma_{18} - 4\widetilde{\sigma}_{12} - 15\widetilde{\sigma}_9 - 15\widetilde{\sigma}_6 - 12\sigma_8 - 28\widetilde{\sigma}_5 - 12\sigma_7 - 21\sigma_6 - 81\sigma_2 - 73\sigma_1 - 51\,,\\
&\rho_7 = 36\sigma_{14} + 41\sigma_{13} + 33\widetilde{\sigma}_{10} + 6\sigma_{12} + 12\sigma_{11} + 25\widetilde{\sigma}_8 + 6\sigma_{10} + 25\sigma_9 + 258\widetilde{\sigma}_4 + 50\widetilde{\sigma}_3 + 55\sigma_3 + 353\widetilde{\sigma}_1\\
&\qquad -30\sigma_{18} - 12\widetilde{\sigma}_{12} - \sigma_{17} - 43\widetilde{\sigma}_9 - 45\widetilde{\sigma}_6 - 29\sigma_8 - 100\widetilde{\sigma}_5 - 32\sigma_7 - 86\sigma_6 - 2\sigma_5 - 226\sigma_2 - 256\sigma_1 - 182\,,\\
&\overline{\rho_7} = \sigma_{17} + \sigma_{14} + 41\sigma_{13} + 39\sigma_{12} + 25\widetilde{\sigma}_9 + 12\sigma_{11} + 23\widetilde{\sigma}_8 + 39\sigma_{10} + 27\sigma_9 + 2\sigma_5 + 307\widetilde{\sigma}_4 + 40\widetilde{\sigma}_3 + 181\sigma_3\\
&\qquad +343\widetilde{\sigma}_1 - 28\sigma_{18} - 12\widetilde{\sigma}_{12} - 41\widetilde{\sigma}_9 - 47\widetilde{\sigma}_6 - 52\sigma_8 - 31\widetilde{\sigma}_5 - 61\sigma_7 - 49\sigma_6 - 278\sigma_2 - 155\sigma_1 - 85\,,\\
&\rho_8 + \overline{\rho_8} = 5\sigma_{14} + 6\sigma_{13} + 4\widetilde{\sigma}_{10} + 2\sigma_{12} + 2\sigma_{11} + 2\widetilde{\sigma}_8 + 3\sigma_{10} + 4\sigma_9 + \sigma_5 + 44\widetilde{\sigma}_4 + 7\widetilde{\sigma}_3 + 12\sigma_3 + 66\widetilde{\sigma}_1\\
&\qquad -4\sigma_8 - 4\sigma_{18} - 2\widetilde{\sigma}_{12} - 6\widetilde{\sigma}_9 - 7\widetilde{\sigma}_6 - 16\widetilde{\sigma}_5 - 6\sigma_7 - 13\sigma_6 - 44\sigma_2 - 46\sigma_1 - 33\,,\\
&\rho_9 + \rho_{15} = 12\sigma_{18} + 5\widetilde{\sigma}_{12} + 17\widetilde{\sigma}_9 + 20\widetilde{\sigma}_6 + 22\sigma_8 + 18\widetilde{\sigma}_5 + 25\sigma_7 + 25\sigma_6 + 125\sigma_2 + 82\sigma_1 + 51\\
&\qquad -3\sigma_{14} - 17\sigma_{13} - 12\widetilde{\sigma}_{10} - 14\sigma_{12} - 5\sigma_{11} - 10\widetilde{\sigma}_8 - 14\sigma_{10} - 11\sigma_9 - 133\widetilde{\sigma}_4 - 23\widetilde{\sigma}_3 - 69\sigma_3 - 159\widetilde{\sigma}_1\,,\\
&\rho_{10} + \overline{\rho_{10}} = 22\sigma_{18} + 10\widetilde{\sigma}_{12} + 31\widetilde{\sigma}_9 + 36\widetilde{\sigma}_6 + 53\sigma_8 + 24\widetilde{\sigma}_5 + 55\sigma_7 + 44\sigma_6 + \sigma_5 + 270\sigma_2 + 150\sigma_1 + 87\\
&\qquad -32\sigma_{13} - 22\widetilde{\sigma}_{10} - 33\sigma_{12} - 10\sigma_{11} - 17\widetilde{\sigma}_8 - 34\sigma_{10} - 20\sigma_9 - 278\widetilde{\sigma}_4 - 50\widetilde{\sigma}_3 - 161\sigma_3 - 321\widetilde{\sigma}_1\,,\\
&\rho_{11} = 12\sigma_{18} + 5\widetilde{\sigma}_{12} + 17\widetilde{\sigma}_9 + 19\widetilde{\sigma}_6 + 19\sigma_8 + 28\widetilde{\sigma}_5 + 20\sigma_7 + \sigma_5 + 30\sigma_6 + 116\sigma_2 + 94\sigma_1 + 63\\
&\qquad -8\sigma_{14} - 17\sigma_{13} - 12\widetilde{\sigma}_{10} - 9\sigma_{12} - 10\widetilde{\sigma}_8 - 5\sigma_{11} - 9\sigma_{10} - 11\sigma_9 - 125\widetilde{\sigma}_4 - 23\widetilde{\sigma}_3 - 51\sigma_3 - 158\widetilde{\sigma}_1\,,\\
&\rho_{12} = \sigma_3\,,\\
&\rho_{13} = 14\sigma_{13} + 10\widetilde{\sigma}_{10} + 14\sigma_{12} + 4\sigma_{11} + 8\widetilde{\sigma}_8 + 14\sigma_{10} + 9\sigma_9 + 110\widetilde{\sigma}_4 + 18\widetilde{\sigma}_3 + 65\sigma_3 + 124\widetilde{\sigma}_1\\
&\qquad -10\sigma_{18} - 4\widetilde{\sigma}_{12} - 14\widetilde{\sigma}_9 - 16\widetilde{\sigma}_6 - 21\sigma_8 - 10\widetilde{\sigma}_5 - 22\sigma_7 - 16\sigma_6 - \sigma_5 - 102\sigma_2 - 53\sigma_1 - 30\,,\\
&\rho_{14} + \rho_{15} = 12\sigma_{18} + 5\widetilde{\sigma}_{12} + 18\widetilde{\sigma}_9 + 19\widetilde{\sigma}_6 + 21\sigma_8 + 16\widetilde{\sigma}_5 + 23\sigma_7 + 21\sigma_6 + 111\sigma_2 + 65\sigma_1 + 38\\
&\qquad -2\sigma_{14} - 17\sigma_{13} - 12\widetilde{\sigma}_{10} - 15\sigma_{12} - 5\sigma_{11} - 10\widetilde{\sigma}_8 - 15\sigma_{10} - 11\sigma_9 - 124\widetilde{\sigma}_4 - 17\widetilde{\sigma}_3 - 69\sigma_3 - 141\widetilde{\sigma}_1\,,
\end{aligned}$$

$$\begin{aligned}
2\rho_{15} &= 30\sigma_{14}+39\sigma_{13}+28\widetilde{\sigma}_{10}+11\sigma_{12}+12\sigma_{11}+22\widetilde{\sigma}_8+11\sigma_{10}+25\sigma_9+244\widetilde{\sigma}_4+40\widetilde{\sigma}_3+67\sigma_3\\
&\quad +326\widetilde{\sigma}_1-28\sigma_{18}-12\widetilde{\sigma}_{12}-40\widetilde{\sigma}_9-44\widetilde{\sigma}_6-26\sigma_8-90\widetilde{\sigma}_5-33\sigma_7-76\sigma_6-211\sigma_2-229\sigma_1-158\,,\\
\rho_{16} &= \widetilde{\sigma}_4+4\widetilde{\sigma}_1-2\sigma_2-2\sigma_1-2\,,\\
\rho_{17} &= 5\sigma_{18}+2\widetilde{\sigma}_{12}+7\widetilde{\sigma}_9+7\widetilde{\sigma}_6+12\sigma_8+5\widetilde{\sigma}_5+11\sigma_7+8\sigma_6+\sigma_5+54\sigma_2+27\sigma_1+15\\
&\quad -7\sigma_{13}-5\widetilde{\sigma}_{10}-7\sigma_{12}-2\sigma_{11}-4\widetilde{\sigma}_8-7\sigma_{10}-4\sigma_9-57\widetilde{\sigma}_4-10\widetilde{\sigma}_3-34\sigma_3-64\widetilde{\sigma}_1\,,\\
\rho_{18} &= \sigma_{14}+6\sigma_{13}+4\widetilde{\sigma}_{10}+6\sigma_{12}+2\sigma_{11}+2\widetilde{\sigma}_8+6\sigma_{10}+4\sigma_9+46\widetilde{\sigma}_4+7\widetilde{\sigma}_3+24\sigma_3+58\widetilde{\sigma}_1\\
&\quad -4\sigma_{18}-2\widetilde{\sigma}_{12}-6\widetilde{\sigma}_9-7\widetilde{\sigma}_6-7\sigma_8-8\widetilde{\sigma}_5-8\sigma_7-8\sigma_6-44\sigma_2-30\sigma_1-19\,,\\
\rho_{19} &= 8\sigma_{14}+10\sigma_{13}+7\widetilde{\sigma}_{10}+2\sigma_{12}+3\sigma_{11}+6\widetilde{\sigma}_8+2\sigma_{10}+6\sigma_9+68\widetilde{\sigma}_4+14\widetilde{\sigma}_3+18\sigma_3+95\widetilde{\sigma}_1\\
&\quad -7\sigma_{18}-3\widetilde{\sigma}_{12}-10\widetilde{\sigma}_9-11\widetilde{\sigma}_6-8\sigma_8-23\widetilde{\sigma}_5-9\sigma_7-21\sigma_6-64\sigma_2-67\sigma_1-48\,,\\
\rho_{20} &= 41\sigma_{13}+29\widetilde{\sigma}_{10}+41\sigma_{12}+12\sigma_{11}+24\widetilde{\sigma}_8+41\sigma_{10}+26\sigma_9+318\widetilde{\sigma}_4+45\widetilde{\sigma}_3+191\sigma_3+346\widetilde{\sigma}_1\\
&\quad -29\sigma_{18}-12\widetilde{\sigma}_{12}-42\widetilde{\sigma}_9-46\widetilde{\sigma}_6-59\sigma_8-29\widetilde{\sigma}_5-64\sigma_7-49\sigma_6-287\sigma_2-149\sigma_1-78\,.
\end{aligned}$$

We considered more Co$_0$-irreps than M_{24} ones; the additional relations between restrictions are:

$$\begin{aligned}
&\widetilde{\sigma}_2+\widetilde{\sigma}_1=\sigma_3+\sigma_1+1\,,\\
&\widetilde{\sigma}_5+\widetilde{\sigma}_4+\sigma_3+1=\sigma_7+\sigma_4+\sigma_2\,,\\
&\widetilde{\sigma}_{11}+\widetilde{\sigma}_3=\widetilde{\sigma}_8+\widetilde{\sigma}_5+\sigma_7+\sigma_4+2\widetilde{\sigma}_2+\sigma_2+\sigma_1+\widetilde{\sigma}_1\,,\\
&\sigma_{14}+\sigma_8=\widetilde{\sigma}_8+\widetilde{\sigma}_7+\sigma_7+2\sigma_4+\widetilde{\sigma}_2+\sigma_2+2\sigma_1\,,\\
&\widetilde{\sigma}_8+\widetilde{\sigma}_7+\sigma_2=\sigma_{12}+\sigma_{10}+\sigma_6+\sigma_3+\widetilde{\sigma}_1\,,\\
&\sigma_{19}+\sigma_8=\sigma_{18}+2\sigma_7+2\sigma_4+\sigma_2+2\sigma_1\,,\\
&\widetilde{\sigma}_6+4\widetilde{\sigma}_4+3\widetilde{\sigma}_3+24535\widetilde{\sigma}_2+24541\widetilde{\sigma}_1=2\sigma_8+2\sigma_5+24533\sigma_3+6\sigma_2+24537\sigma_1+24537\,,\\
&85\sigma_{14}+88\sigma_{13}+62\widetilde{\sigma}_{10}+4\sigma_{12}+26\sigma_{11}+50\widetilde{\sigma}_8+4\sigma_{10}+56\sigma_9+516\widetilde{\sigma}_4+97\widetilde{\sigma}_3+72\sigma_3+754\widetilde{\sigma}_1\\
&\qquad =62\sigma_{18}+26\widetilde{\sigma}_{12}+90\widetilde{\sigma}_9+99\widetilde{\sigma}_6+41\sigma_8+234\widetilde{\sigma}_5+53\sigma_7+190\sigma_6+453\sigma_2+582\sigma_1+428\,.
\end{aligned}$$

Of course we'd prefer true representations to virtual ones, but perhaps we should not be too surprised that virtual representations arise here, because the definition (2.1) of elliptic genus involves signs, and after all even H_{00} and H_0 were virtual (but see Section 6).

However, the dimensions of the virtual Co$_0$-representations needed will be *extremely* large. By the *total dimension* of a virtual representation $\rho_+ \ominus \rho_-$ we mean the quantity $\dim(\rho_+)+\dim(\rho_-)$. For any M_{24} representation, if there is any Co$_0$-virtual representation which restricts to it, there will be infinitely many (just add to it any of the above relations). For each basic combination of M_{24}-irreps, we have selected above the Co$_0$-virtual representation of smallest total dimension we could find, which restricts to it. (This ambiguity would be eliminated by identifying if possible the twining elliptic genera ϕ_g for all $g \in \mathrm{Co}_0$.)

For example, of all Co$_0$-irreps of dimension less than 1 million, only 2 of them contain ρ_2 or $\overline{\rho_2}$ as a summand: namely, a Co$_1$-irrep of dim 313950, and a spinor irrep of dim 789360. The smallest virtual Co$_0$-representation we could find which restricts to $\rho_2 + \overline{\rho_2}$ (a 90-dimensional representation of M_{24}) has total dimension over 100 billion.

In fact more is true. Suppose for contradiction the natural identification of the Witten index w_g with $\widetilde{\sigma}_1$ (the only Co_0-rep which restricts to $1 + \rho_1$). Now, [30] compute (among other things) the twining characters for the automorphism of the Gepner model $(1)^6$; they found two symmetries of order 9 (the bottom two rows of their Table 1), both of whose Witten indices equal 3, and whose twining elliptic genera are nonetheless different. However, Co_0 has only one conjugacy class (namely class 9C) of order 9 with $\widetilde{\sigma}_1 = 3$, and this would imply those twining genera should be equal. (We thank Roberto Volpato for sharing this observation.) This contradiction means that w_g cannot equal $\widetilde{\sigma}_1(g)$, and so the Witten genus too will be virtual, having total dimension much greater than 24. These large total dimensions make the proposed extension to the Conway groups seem unlikely.

6 Speculations

Because the phenomena underlying Mathieu Moonshine are still obscure, we conclude with assorted questions and speculations. The most obvious challenge suggested by this paper is to explicitly construct these representations H_n. The evidence that there is some vertex algebra-like object underlying the Mathieu Moonshine observations now seems overwhelming; perhaps the most satisfying way to construct the H_n would be to construct this vertex (super)algebra. This would provide the algebraic underpinning of Mathieu Moonshine, and it should hint at its still-mysterious geometric and physical meanings.

We regard this vertex superalgebra construction as the most important challenge of Mathieu Moonshine. At $c = 24$ we have an $N = 0$ VOA (namely the Moonshine Module $V^\natural$) with lots of nice properties including an action of $\mathbb{M}$. At $c = 12$ we have an $N = 1$ VOSA (namely Duncan's algebra [16]), with lots of nice properties including an action of the Conway group. Could there be at $c = 6$ an $N = 2$ or $N = 4$ VOSA (namely the algebra underlying our Mathieu Moonshine) with lots of nice properties including an action of M_{24}? After all, M_{24}, Co_1, $\mathbb{M}$ is the Holy Trinity of sporadic finite simple groups (e.g. Griess [35] constructed the Monster by starting with M_{24}, lifting to Conway and then moving on to the Monster). (I thank Gerald Höhn and Chongying Dong for informal discussions on this point).

Another possible construction, which may also bring K3 into Mathieu Moonshine, is through the chiral de Rham complex [40] (a sheaf of vertex superalgebras) associated to a K3 surface. The trace over its global section (which is itself a vertex superalgebra) recovers elliptic genus [3]. The orbifold theory, including the construction of twisted modules, has also been studied [26].

One of the most intriguing aspects of Mathieu Moonshine is the *positivity* of the irrep multiplicities. This discussion is already anticipated in Chapter 7.2 of Wendland's thesis [52]. There we find that the partition function in the Ramond-

Ramond sector of an $N = 4$ $c = 6$ superconformal field theory is

$$\begin{aligned} Z_{RR} =&|ch^s_{0,0}|^2 + h|ch^s_{1/4,0}|^2 + \epsilon\,(ch^s_{0,0}ch^{s*}_{1/4,0} + c.c.) + F ch^l_{1/4,1/2}ch^{s*}_{0,0} + F' ch^s_{0,0}ch^{l*}_{1/4,1/2} \\ &+ G ch^l_{1/4,1/2}ch^{s*}_{1/4,0} + G' ch^s_{1/4,0}ch^{l*}_{1/4,1/2} + H|ch^l_{1/4,1/2}|^2 \,, \end{aligned} \tag{6.1}$$

using the same conventions for $N = 4$ superconformal characters as (2.6), where h, ϵ are nonnegative integers and F, F', G, G', H are functions of τ, z with only nonnegative integer coefficients in $q, \overline{q}$. The elliptic genus is

$$\phi = (\epsilon - 2)ch^s_{0,0} + (h - 2\epsilon)ch^s_{1/4,0} + (G - 2F)ch^l_{1/4,1/2} \,. \tag{6.2}$$

For the K3 component of moduli space, $\epsilon = 0$ and $h = h^{1,1} = 20$. In particular note that the function F counts holomorphic fields which exist in that theory but which are not contained in the $N = 4$ vacuum representation. We can think of holomorphic fields, very roughly, as additional symmetries of the theory; since we would expect that generically the chiral algebra of the K3 sigma model is simply $N = 4$ superconformal, we should have generically $F = 0$. This would mean the non-BPS part of the elliptic genus is (generically) this function G which has only non-negative coefficients. From this the conjectured non-negativity would follow. This argument also applies to the character-valued elliptic genus, so all multiplicities should be nonnegative. A similar argument can be found in Ooguri [44]. (We thank Katrin Wendland for discussions on this point.)

It is commonly expected that the moduli space of $c = 6$ $N = 4$ superconformal field theories consists of two components: a toroidal one with vanishing elliptic genus, and the K3 sigma models with elliptic genus $2\phi_{0,1}$. But if this is true, then there can be no $c = 6$ $N = 4$ superconformal field theory underlying Mathieu moonshine: the work of [30] shows that the corresponding automorphism groups are too small.

But $N = 4$ (more precisely, $N = (4, 4)$) theories possess more supersymmetry than we need. We could deform an $N = (4, 4)$ theory to e.g. an $N = (0, 4)$ heterotic (see e.g. [55] for an introduction to similar theories). These theories are geometrical, corresponding to bundles over K3, and are also related to chiral de Rham. They possess the desired elliptic genus $2\phi_{0,1}$. The moduli space of these theories is 90-dimensional, far larger than that of $N = (4, 4)$, so there is a much greater chance for some larger symmetry groups. However, the breaking of the left-moving $N = 4$ supersymmetry seems to destroy the justification for decomposing the elliptic genus into $N = 4$ superconformal characters — a step crucial to Mathieu moonshine. It is tempting to partially break the supersymmetry, say to $N = (4, 1)$, but such theories seem to possess the full $N = (4, 4)$ supersymmetry. (I thank Ilarion Melnikov for discussions involving this point.)

The difficulty in interpreting Mathieu Moonshine in terms of K3 sigma models (e.g. the interesting work of Taormina-Wendland [50] is still far from realising M_{24}) leads one to consider the unhappy possibility that the connection with K3 is perhaps accidental. The relevant (twisted) elliptic genera are so heavily constrained that there are bound to be empty coincidences. For example, any of the 71 or so $c = 24$

holomorphic rational VOA — e.g. one associated to the Leech lattice — will have a VOA character (a.k.a. partition function) equal to $J(\tau)+c$ for some constant $c \in \mathbb{Z}_{\geq 0}$, again because it is so severely constrained. The coefficients of J (with or without c) will have an interpretation as dimensions of Monster representations as we know, but conjecturally only one of those VOAs actually carries a nontrivial action of the Monster $\mathbb{M}$. It takes a (slightly) deeper analysis to rule out Monster actions on these other VOAs. Could this relation of K3 sigma models to our Mathieu Moonshine be likewise illusory? After all, it is clear that the Jacobi forms of Umbral Moonshine cannot have a direct interpretation as elliptic genera, when the group is not M_{24}.

Similarly, perhaps we shouldn't regard M_{24} as sacrosanct. The evenness property of Theorem B hints perhaps that the symmetry is somewhat larger. The analysis of [30] and our Theorem C hints that the 'ultimate' symmetry lies somewhere between M_{24} and Conway. The split extension $\mathbb{Z}_2^{12} \rtimes M_{24}$, a maximal subgroup of Co_0, is the only such group and seems worth a look. (The evenness property was used in [12] to prove one of the Umbral Moonshine conjectures, and together with Theorem A proves that the elliptic genus of Enriques surfaces decomposes into a sum of M_{12} characters.)

The appearance of $\Gamma_0(|g|)$ in Lemma 1 is what one would expect from the CFT orbifold story. Let us quickly review that basic theory, as developed by [14],[15],[1], and others. Let $\mathcal{V}$ be a (bosonic) rational VOA with automorphisms containing some finite group G. By $\mathcal{V}^G$ we mean the vertex operator subalgebra consisting of all fixed-points of G in $\mathcal{V}$. Conjecturally, $\mathcal{V}^G$ is also rational; in the mathematical literature it is called the *orbifold* of $\mathcal{V}$ by G (the orbifold construction means something a little different in the physics literature).

We are interested here in the simplest case, where $\mathcal{V}$ only has a single irreducible module (namely itself). The twisted modules $M(g)$ of $\mathcal{V}$ are parametrised by (a subset of the) conjugacy class representatives g, and (a subgroup of) the centraliser of g in G acts on them, so we can define *twisted twining characters* $Z_{g,h}(\tau) = \mathrm{Tr}_{M_g}(hq^{L_0-c/24})$. The special case $Z_{1,h}(\tau)$ are sometimes called *McKay-Thompson series* in the mathematics literature. Then for any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, $Z_{g,h}\left(\frac{a\tau+b}{c\tau+d}\right)$ will equal $Z_{g^a h^c, g^b h^d}(\tau)$ up to some phase. For general g, h, $Z_{g,h}(\tau)$ will thus be a modular function with multiplier for the group of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(\mathrm{lcm}(|g|,|h|))$. However, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(|h|)$ will send the McKay-Thompson series $Z_{1,h}$ to another, Z_{1,h^d} (up to a phase). The integer d will be coprime to $|h|$, and sends a character value of h to a Galois associate. When the McKay-Thompson series have integer coefficients (e.g. in Monstrous or Mathieu Moonshine), these Galois associates have the same value, and $Z_{1,h^d} = Z_{1,h}$. In other words, since we have that integrality, our McKay-Thompson series $Z_{1,h}$ will be modular functions (with multiplier) for $\Gamma_0(|h|)$.

To complete the story, we need to consider equivariant cohomology. Let B_G be a classifying space, then the group cohomology $H^*_G(1;N)$ (often denoted $H^*(G;N)$) with values in a G-module N is defined by $H^*(B_G;N)$. Just as $H^2_G(1;U(1))$ controls the projective representatives of G, and $\alpha \in H^3_G(G;\mathbb{Z})$ (G acting on itself by conjugation) parametrises the possible orbifold fusion rings (all given by twisted equivariant K-theory ${}^{\alpha}K_G^0(G)$), the representation theory of $\mathcal{V}^G$ is controlled by a 3-cocycle $\alpha \in H_G^3(1;U(1))$. In particular, when α is trivial, the modules M of $\mathcal{V}^G$ conjecturally are in natural one-to-one correspondence with pairs (g,ψ), where g is a conjugacy class representative in G and ψ is an irrep of the centraliser $C_G(g)$. Their characters $\chi_M(\tau) = \mathrm{Tr}_M q^{L_0-c/24}$ form a vector-valued modular function for $\mathrm{SL}_2(\mathbb{Z})$ with multiplier explicitly defined in terms of G [14]. The twisted modules $M(g)$ of $\mathcal{V}$ are also (conjecturally) parametrised by all conjugacy class representatives g, and the full centraliser of g in G acts on them Then $Z_{g^k,h^k}(\tau) = Z_{g,h}(\tau)$ and $Z_{g,h}\left(\frac{a\tau+b}{c\tau+d}\right) = Z_{g^ah^c,g^bh^d}(\tau)$. When the cocycle is not trivial — the generic case — a subset of these pairs (g,ψ) and (g,h) will not be viable. Again, the characters of $\mathcal{V}^G$ should yield a vector-valued modular form, but the multiplier is more complicated (though known [11]). Similarly, the relations $Z_{g^k,h^k}(\tau) = Z_{g,h}(\tau)$ and $Z_{g,h}\left(\frac{a\tau+b}{c\tau+d}\right) = Z_{g^ah^c,g^bh^d}(\tau)$ have to be adjusted by phases determined by α.

We have order-12 phases in Mathieu Moonshine (e.g. the $e^{2\pi i cd/(|g|h)}$ in (2.17)), so we would expect (supposing the orbifold theory extends to our $N=4$ setting) the relevant 3-cocycle α to have order 12. Indeed, [18] computed that $H^3_{M24}(1;U(1)) \cong \mathbb{Z}_{12}$, and [31] explicitly verified that a generating cocycle yields the phases appearing in Mathieu Moonshine. Moreover, $H^3_{M23}(1;U(1)) \cong 1$ [41], and so the phases for elements lying in M_{23} should be trivial, and this indeed is what is observed. It is also observed that many pairs (g,h) are not viable, as expected since α is nontrivial. If some group G larger than M_{24} (e.g. Co_0 or $2^{12}.M_{24}$) actually acts here, then for this reason we'd require $H^3_G(1;U(1))$ to contain an order-24 element. (The cohomology of Co_0 has not been computed yet). It is observed in [31] that some twisted twining elliptic genera vanish even when there is no cohomological obstruction — this would be surprising for a bosonic theory but elliptic genera (being a signed trace) often vanishes in nontrivial theories so this is not mysterious here.

In summary, the perfect formal fit of twisted twining elliptic genera here with the general theory of VOA orbifolds, lends strong support to the belief that there is a vertex operator superalgebra underlying Mathieu Moonshine.

In the case of Monstrous Moonshine, the phases are 24th roots of 1, so we would expect the orbifold $(\mathcal{V}^\natural)^{\mathbb{M}}$ to be governed by an *order 24* cocycle. We would also expect this α to obstruct certain pairs $Z_{g,h}$. Indeed, empirical observations encapsulated in Norton's *Generalised Monstrous Moonshine* [43] make exactly this point: whether a pair is viable or not is captured by Norton's Fricke dichotomy. This should be equivalent to the cohomology condition. This relation between cohomology and Norton's mysterious Fricke condition seems to be new. It would be very interesting to make that relation more explicit.

Acknowledgements. It is my pleasure to thank Matthias Gaberdiel and, more generally, the conference participants of the Mathieu Moonshine workshop at ETH in July 2011, for discussions. I thank Thomas Creutzig, Matthias Gaberdiel, Gerald Höhn, Ilarion Melnikov, Roberto Volpato and Katrin Wendland for comments involving an earlier draft of this manuscript.

References

[1] P. Bantay: *Orbifolds, Hopf algebras and the Moonshine.* Lett. Math. Phys. 22 (1991), 187–194.

[2] R. E. Borcherds: *Monstrous moonshine and monstrous Lie superalgebras.* Invent. Math. 109 (1992), 405–444.

[3] L. A. Borisov, A. Libgober: *Elliptic genera of toric varieties and applications to mirror symmetry.* Invent. Math. 140 (2000), 453–485.

[4] K. Bringmann, K. Ono: *The $f(q)$ mock theta function conjecture and partition ranks.* Invent. Math. 165 (2006), 243–266.

[5] M.C.N. Cheng: *K3 Surfaces, N=4 dyons, and the Mathieu group M24.* Commun. Number Theory Phys. 4 (2010), 623 [arXiv:1005.5415].

[6] M.C.N. Cheng, J. F. R. Duncan: *On Rademacher sums, the largest Mathieu group, and the holographic modularity of Moonshine.* [arXiv:1110.3859].

[7] M. C. N. Cheng, J. F. R. Duncan, J. A. Harvey: *Umbral moonshine.* [arXiv:1204.2779].

[8] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker, R.A. Wilson: *The Atlas of Finite Groups.* (Oxford University Press, Oxford 2003).

[9] J.H. Conway, S. Norton: *Monstrous moonshine.* Bull. Lond. Math. Soc. 11 (1979), 308–339.

[10] J.H. Conway, N. J. A. Sloane: *Sphere Packings, Lattices and Groups.* (3rd edition) (Springer, New York 1999).

[11] A. Coste, T. Gannon, P. Ruelle: *Finite group modular data.* Nucl. Phys. B581 (2000), 679–717.

[12] T. Creutzig, G. Höhn, T. Miezaki: *The McKay-Thompson series of Mathieu Moonshine modulo two.* [arXiv: 1211.3703v2].

[13] A. Dabholkar, S. Murthy, D. Zagier: *Quantum black holes, wall crossing, and mock modular forms.* [arXiv: 1208.4074v1].

[14] R. Dijkgraaf, C. Vafa, E. Verlinde, H. Verlinde: *The operator algebra of orbifold models.* Commun. Math. Phys. 123 (1989), 485–526.

[15] R. Dijkgraaf, E. Witten: *Topological gauge theories and group cohomology.* Commun. Math. Phys. 129 (1990), 393–429.

[16] J. F. Duncan: *Super-Moonshine for Conway's largest sporadic group.* Duke Math. J. 139 (2007), 255–315.

[17] J. F. R. Duncan, I. B. Frenkel: *Rademacher sums, moonshine and gravity.* Commun. Number Theor. Phys. 5 (2011) no. 4, 1–128; [arXiv:0907.4529].

[18] M. Dutour Sikirić, G. Ellis: *Wythoff polytopes and low-dimensional homology of Mathieu groups.* J. Algebra 322 (2009), 4143–4150.

[19] T. Eguchi, K. Hikami: *Superconformal algebras and mock theta functions 2. Rademacher expansion for K3 surface.* Commun. Number Theory Phys. 3 (2009), 531–554; [arXiv: 0904.0911v2]

[20] T. Eguchi, K. Hikami: *Note on twisted elliptic genus of K3 surface.* Phys. Lett. B694 446–455 (2011) [arXiv:1008.4924]

[21] T. Eguchi, K. Hikami: *$N = 2$ moonshine.* Phys. Lett. B717 (2012), 266–273; [arXiv:1209.0610].

[22] T. Eguchi, K. Hikami: *Enriques moonshine.* [arXiv:1301.5043]

[23] T. Eguchi, H. Ooguri, A. Taormina, S.-K. Yang: *Superconformal algebras and string compactification on manifolds with SU(N) holonomy.* Nucl. Phys. B315 (1989), 193–221.

[24] T. Eguchi, H. Ooguri, Y. Tachikawa: *Notes on the K3 surface and the Mathieu group M24.* Exper. Math. 20 91–96 (2011); [arXiv:1004.0956].

[25] M. Eichler, D. Zagier: *The Theory of Jacobi Forms.* (Birkhäuser, Boston, 1985).

[26] E. Frenkel, M. Szczesny: *Chiral de Rham complex and orbifolds.* J. Alg. Geom. 16 (2007), 599–624.

[27] I. Frenkel, J. Lepowsky, A. Meurman: *Vertex Operator Algebras and the Monster* (Academic Press, San Diego 1988).

[28] M. R. Gaberdiel, S. Hohenegger, R. Volpato: *Mathieu twining characters for K3.* JHEP 09 (2010) 058; [arXiv:1006.0221].

[29] M. R. Gaberdiel, S. Hohenegger, R. Volpato: *Mathieu Moonshine in the elliptic genus of K3.* JHEP 10 (2010) 062; [arXiv:1008.3778v3].

[30] M. R. Gaberdiel, S. Hohenegger, R. Volpato: *Symmetries of K3 sigma models.* [arXiv:1106.4315v1].

[31] M. R. Gaberdiel, D. Persson, H. Ronellenfitsch, R. Volpato: *Generalised Mathieu Moonshine.* [arXiv: 1211.7074].

[32] T. Gannon: *Moonshine beyond the Monster: The Bridge connecting Algebra, Modular Forms and Physics.* (Cambridge University Press, Cambridge 2006).

[33] D. Goldfeld, P. Sarnak: *Sums of Kloosterman sums.* Invent. Math. 71 (1983), 243–250.

[34] S. Govindarajan: *Unravelling Mathieu moonshine.* [arXiv: 1106.5715v1].

[35] R. L. Griess, Jr.: *The friendly giant.* Invent. Math. 69 (1982), 1–102.

[36] C. Hooley: *An asymptotic formulae in the theory of numbers.* Proc. London Math. Soc. 7 (1957), 396–413.

[37] C. Hooley: *On the number of divisors of quadratic polynomials.* Acta Math. 110 (1963), 97–114.

[38] Hua Loo Keng: *Introduction to Number Theory.* (Springer, Berlin 1982).

[39] S. Kondo: *Niemeier lattices, Mathieu groups and finite groups of symplectic automorphisms of K3 surfaces.* Duke Math. J. 92 (1998), 593–603, appendix by S. Mukai.

[40] F. Malikov, V. Schechtman, A. Vaintrob: *Chiral de Rham complex.* Commun. Math. Phys. 204 (1999), 439–473.

[41] R.J. Milgram: *The cohomology of the Mathieu group M23.* J. Group Theory 3 (2000), 7–26.

[42] S. Mukai: *Finite groups of automorphisms of K3 surfaces and the Mathieu group.* Invent. Math. 94 (1988), 183–221.

[43] S. P. Norton: *Generalized moonshine.* The Arcata Conference on Representations of Finite Groups (Arcata, 1986) Proc. Sympos. Pure Math. 47 (American Mathematical Society, Providence 1987) 208–209.

[44] H. Ooguri: *Superconformal symmetry and geometry of Ricci flat Kahler manifolds.* Int. J. Mod. Phys. A4 (1989), 4304–4324.

[45] H. Rademacher: *Bestimmung einer gewissen Einheitswurzel in der Theorie der Modulfunktionen.* J. London Math. Soc. 7 (1932), 14–19.

[46] H. Rademacher: *On the Selberg formula for* $A_k(n)$. J. Indian Math. Soc. 21 (1957), 41–55.

[47] A. Selberg: *On the estimation of Fourier coefficients of modular forms.* in: *Proc. Sympos. Pure Math., Vol. VIII* (American Mathematical Society, Providence 1965), pp. 1–15.

[48] S.D. Smith: *On the head characters of the monster simple group.* Finite Groups – Coming of Age (Montreal 1982) (American Mathematical Society, Providence 1985) pp. 303–313.

[49] J. Sturm: *On the congruence of modular forms.* Number theory (New York, 1984–1985) *Lecture Notes in Math.* 1240 (Springer, Berlin, 1987), 275–280.

[50] A. Taormina, K. Wendland: *The overarching finite symmetry group of Kummer surfaces in the Mathieu group* M_{24}*;* [arXiv:1107.3834].

[51] A. L. Whiteman: *A sum connected with the series for the partition function.* Pacific J. Math. 6 (1956), 159–176.

[52] K. Wendland: *Moduli spaces of unitary conformal field theories.* Ph.D. Thesis, Universität Bonn (2000).

[53] R. A. Wilson: *On the 3-local subgroups of Conway's group* Co_1. J. Alg. 113 (1988), 261–262.

[54] E. Witten: *Elliptic genera and quantum field theory.* Commun. Math. Phys. 109 (1987), 525–536.

[55] E. Witten: *Two-dimensional models with (0,2) supersymmetry: perturbative aspects.* Adv. Theor. Math. Phys. 11 (2007), 1–63.

What is moonshine?

Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998). Doc. Math. 1998, Extra Vol. I, 607–615

Richard E. Borcherds, [*]
D.P.M.M.S., 16 Mill Lane, Cambridge, CB2 1SB, England.
e-mail: reb@dpmms.cam.ac.uk
home page: www.dpmms.cam.ac.uk/˜reb

This is an informal write up of my talk at the I.C.M. in Berlin. It gives some background to Goddard's talk [Go] about the moonshine conjectures. For other survey talks about similar topics see [B94], [B98], [LZ], [J], [Ge], [Y].

The classification of finite simple groups shows that every finite simple group either fits into one of about 20 infinite families, or is one of 26 exceptions, called sporadic simple groups. The monster simple group is the largest of the sporadic finite simple groups, and was discovered by Fischer and Griess [G]. Its order is

$$\begin{aligned}&8080, 17424, 79451, 28758, 86459, 90496, 17107, 57005, 75436, 80000, 00000\\ =&2^{46}.3^{20}.5^{9}.7^{6}.11^{2}.13^{3}.17.19.23.29.31.41.47.59.71\end{aligned}$$

(which is roughly the number of elementary particles in the earth). The smallest irreducible representations have dimensions $1, 196883, 21296876, \ldots$. The elliptic modular function $j(\tau)$ has the power series expansion

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + \ldots$$

where $q = e^{2\pi i\tau}$, and is in some sense the simplest nonconstant function satisfying the functional equations $j(\tau) = j(\tau + 1) = j(-1/\tau)$. John McKay noticed some rather weird relations between coefficients of the elliptic modular function and the representations of the monster as follows:

$$\begin{aligned}1 &= 1\\ 196884 &= 196883 + 1\\ 21493760 &= 21296876 + 196883 + 1\end{aligned}$$

where the numbers on the left are coefficients of $j(\tau)$ and the numbers on the right are dimensions of irreducible representations of the monster. At the time he discovered these relations, several people thought it so unlikely that there could be a relation between the monster and the elliptic modular function that they politely told McKay that he was talking nonsense. The term “monstrous moonshine” (coined by Conway) refers to various extensions of McKay's observation, and in particular to relations between sporadic simple groups and modular functions.

For the benefit of readers who are not native English speakers, I had better point out that “moonshine” is not a poetic terms referring to light from the moon. It means

---

[*] Supported by a Royal Society professorship.

foolish or crazy ideas. (Quatsch in German.) A typical example of its use is the following quotation from E. Rutherford (the discoverer of the nucleus of the atom): "The energy produced by the breaking down of the atom is a very poor kind of thing. Anyone who expects a source of power from the transformations of these atoms is talking moonshine." (Moonshine is also a name for corn whiskey, especially if it has been smuggled or distilled illegally.)

We recall the definition of the elliptic modular function $j(\tau)$. The group $SL_2(\mathbf{Z})$ acts on the upper half plane H by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(\tau) = \frac{a\tau + b}{c\tau + d}.$$

A modular function (of level 1) is a function f on H such that $f((a\tau+b)/(c\tau+d)) = f(\tau)$ for all $\left(\begin{smallmatrix} ab \\ cd \end{smallmatrix}\right) \in SL_2(\mathbf{Z})$. It is sufficient to assume that f is invariant under the generators $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$ of $SL_2(\mathbf{Z})$. The elliptic modular function j is the simplest nonconstant example, in the sense that any other modular function can be written as a function of j. It can be defined as follows:

$$\begin{aligned} j(\tau) &= \frac{E_4(\tau)^3}{\Delta(\tau)} \\ &= q^{-1} + 744 + 196884q + 21493760q^2 + \cdots \\ E_4(\tau) &= 1 + 240\sum_{n>0}\sigma_3(n)q^n \qquad\qquad (\sigma_3(n) = \sum_{d|n} d^3) \\ &= 1 + 240q + 2160q^2 + \cdots \\ \Delta(\tau) &= q\prod_{n>0}(1-q^n)^{24} \\ &= q - 24q + 252q^2 + \cdots \end{aligned}$$

A modular form of weight k is a holomorphic function $f(\tau) = \sum_{n\geq 0} c(n)q^n$ on the upper half plane satisfying the functional equation $f((a\tau+b)/(c\tau+d)) = (c\tau+d)^k f(\tau)$ for all $\left(\begin{smallmatrix} ab \\ cd \end{smallmatrix}\right) \in SL_2(\mathbf{Z})$. The function $E_4(\tau)$ is an Eisenstein series and is a modular form of weight 4, while $\Delta(\tau)$ is a modular form of weight 12.

The function $j(\tau)$ is an isomorphism from the quotient $SL_2(\mathbf{Z})\backslash H$ to $\mathbf{C}$, and is uniquely defined by this up to multiplication by a constant or addition of a constant. In particular any other modular function is a function of j, so j is in some sense the simplest nonconstant modular function.

An amusing property of j (which so far seems to have no relation with moonshine) is that $j(\tau)$ is an algebraic integer whenever τ is an imaginary quadratic irrational number. A well known consequence of this is that

$$\exp(\pi\sqrt{163}) = 262537412640768743.99999999999925\ldots$$

is very nearly an integer. The explanation of this is that $j((1+i\sqrt{163})/2)$ is exactly the integer $-262537412640768000 = -2^{18}3^3 5^3 23^3 29^3$, and

$$\begin{aligned} j((1+i\sqrt{163})/2) &= q^{-1} + 744 + 196884q + \cdots \\ &= -e^{\pi\sqrt{163}} + 744 + (\text{something very small}). \end{aligned}$$

McKay and Thompson suggested that there should be a graded representation $V = \oplus_{n\in\mathbf{Z}} V_n$ of the monster, such that $\dim(V_n) = c(n-1)$, where $j(\tau) - 744 = \sum_n c(n)q^n = q^{-1} + 196884q + \cdots$. Obviously this is a vacuous statement if interpreted literally, as we could for example just take each V_n to be a trivial representation. To characterize V, Thompson suggested looking at the McKay-Thompson series

$$T_g(\tau) = \sum_n Tr(g|V_n)q^{n-1}$$

for each element g of the monster. For example, $T_1(\tau)$ should be the elliptic modular function. Conway and Norton [C-N] calculated the first few terms of each McKay-Thomson series by making a reasonable guess for the decomposition of the first few V_n's into irreducible representations of the monster. They discovered the astonishing fact that all the McKay-Thomson series appeared to be Hauptmoduls for certain genus 0 subgroups of $SL_2(\mathbf{R})$. (A Hauptmodul for a subgroup Γ is an isomorphism from $\Gamma\backslash H$ to $\mathbf{C}$, normalized so that its Fourier series expansion starts off $q^{-1} + O(1)$.)

As an example of some Hauptmoduls of elements of the monster, we will look at the elements of order 2. There are 2 conjugacy classes of elements of order 2, usually called the elements of types $2A$ and $2B$. The corresponding McKay-Thompson series start off

$$T_{2B}(\tau) = q^{-1} + 276q - 2048q^2 + \cdots \qquad \text{Hauptmodul for } \Gamma_0(2)$$
$$T_{2A}(\tau) = q^{-1} + 4372q + 96256q^2 + \cdots \qquad \text{Hauptmodul for } \Gamma_0(2)+$$

The group $\Gamma_0(2)$ is $\{\binom{ab}{cd} \in SL_2(\mathbf{Z}) | c \text{ is even}\}$, and the group $\Gamma_0(2)+$ is the normalizer of $\Gamma_0(2)$ in $SL_2(\mathbf{R})$. Ogg had earlier commented on the fact that the full normalizer $\Gamma_0(p)+$ of $\Gamma_0(p)$ for p prime is a genus 0 group if and only if p is one of the primes 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 41, 47, 59, or 71 dividing the order of the monster.

Conway and Norton's conjectures were soon proved by A. O. L. Atkin, P. Fong, and S. D. Smith. The point is that to prove something is a virtual character of a finite group it is only necessary to prove a finite number of congruences. In the case of the moonshine module V, proving the existence of an infinite dimensional representation of the monster whose McKay-Thompson series are given Hauptmoduls requires checking a finite number of congruences and positivity conditions for modular functions, which can be done by computer.

This does not give an explicit construction of V, or an explanation about why the conjectures are true. Frenkel, Lepowsky, and Meurman managed to find an explicit construction of a monster representation $V = \oplus V_n$, such that $\dim(V_n) = c(n-1)$, and this module had the advantage that it came with some extra algebraic structure preserved by the monster. However it was not obvious that V satisfied the Conway-Norton conjectures. So the main problem in moonshine was to show that the monster modules constructed by Frenkel, Lepowsky and Meurman on the one hand, and by Atkin, Fong, and Smith on the other hand, were in fact the same representation of the monster.

Peter Goddard [Go] has given a description of the proof of this in his talk in this volume, so I will only give a quick sketch of this. The main steps of the proof are as follows:

1. The module V constructed by Frenkel, Lepowsky, and Meurman has an algebraic structure making it into a "vertex algebra". A detailed proof of this is given in [F-L-M].
2. Use the vertex algebra structure on V and the Goddard-Thorn no-ghost theorem [G-T] from string theory to construct a Lie algebra acted on by the monster, called the monster Lie algebra.
3. The monster Lie algebra is a "generalized Kac-Moody algebra" ([K90]); use the (twisted) Weyl-Kac denominator formula to show that $T_g(\tau)$ is a "completely replicable function".
4. Y. Martin [M], C. Cummins, and T. Gannon [C-G] proved several theorems showing that completely replicable functions were modular functions of Hauptmoduls for genus 0 groups. By using these theorems it follows that T_g is a Hauptmodul for a genus 0 subgroup of $SL_2(\mathbf{Z})$, and hence V satisfies the moonshine conjectures. (The original proof used an earlier result by Koike [Ko] showing that the appropriate Hauptmoduls were completely replicable, together with a boring case by case check and the fact that a completely replicable function is characterized by its first few coefficients.)

We will now give a brief description of some of the terms above, starting with vertex algebras. The best reference for finding out more about vertex algebras is Kac's book [K]. In this paragraph we give a rather vague description. Suppose that V is a commutative ring acted on by a group G. We can form expressions like

$$u(x)v(y)w(z)$$

where $u, v, w \in V$ and $x, y, z \in G$, and the action of $x \in G$ on $u \in V$ is denoted rather confusingly by $u(x)$. (This is not a misprint for $x(u)$; the reason for this strange notation is to make the formulas compatible with those in quantum field theory, where u would be a quantum field and x a point of space-time.) For each fixed $u, v, \ldots \in V$, we can think of $u(x)v(y)\cdots$ as a function from G^n to V. We can rewrite the axioms for a commutative ring acted on by G in terms of these functions. We can now think of a vertex algebra roughly as follows: we are given lots of functions from G^n to V satisfying the axioms mentioned above, with the difference that these functions are allowed to have certain sorts of singularities. In other words a vertex algebra is a sort of commutative ring acted on by a group G, except that the multiplication is not defined everywhere but has singularities. In particular we cannot recover an underlying ring by defining the product of u and v to be $u(0)v(0)$, because the function $u(x)v(y)$ might happen to have a singularity at $u = v = 0$.

It is easy to write down examples of vertex algebras: any commutative ring acted on by a group G is an example. (Actually this is not quite correct: for technical reasons we should use a formal group G instead of a group G.) Conversely any vertex algebra "without singularities" can be constructed in this way. Unfortunately there are no easy examples of vertex algebras that are not really commutative rings. One reason for this is that nontrivial vertex algebras must be infinite dimensional; the point is that if a vertex algebra has a nontrivial singularity, then by differentiating it we can make the singularity worse and worse, so we must have an infinite dimensional space of singularities. This is only possible if the vertex algebra is infinite dimensional. However there are plenty of important infinite dimensional examples; see for example Kac's book for a construction of the most important examples, and [FLM] for a construction of the monster vertex algebra.

Next we give a brief description of generalized Kac-Moody algebras. The best way to think of these is as infinite dimensional Lie algebras which have most of the good properties of finite dimensional reductive Lie algebras. Consider a typical finite dimensional reductive Lie algebra G, (for example the Lie algebra $G = M_n(\mathbf{R})$ of $n \times n$ real matrices). This has the following properties:

1. G has an invariant symmetric bilinear form $(,)$ (for example $(a, b) = -Tr(a, b)$).
2. G has a (Cartan) involution ω (for example, $\omega(a) = -a^t$).
3. G is graded as $G = \oplus_{n\in\mathbf{Z}} G_n$ with G_n finite dimensional and with ω acting as -1 on the "Cartan subalgebra" G_0. (For example, we could put the basis element $e_{i,j}$ of $M_n(\mathbf{R})$ in G_{i-j}.)
4. $(a, \omega(a)) > 0$ if $g \in G_n$, $g \neq 0$.

Conversely any Lie algebra satisfying the conditions above is essentially a sum of finite dimensional and affine Lie algebras. Generalized Kac-Moody algebras are defined by the same conditions with one small change: we replace condition 4 by

4'. $(a, \omega(a)) > 0$ if $g \in G_n$, $g \neq 0$ and $n \neq 0$.

This has the effect of allowing an enormous number of new examples, such as all Kac-Moody algebras and the Heisenberg Lie algebra (which behaves like a sort of degenerate affine Lie algebra). Generalized Kac-Moody algebras have many of the properties of finite dimensional semisimple Lie algebras, and in particular they have an analogue of the Weyl character formula for some of their representations, and an analogue of the Weyl denominator formula. An example of the Weyl-Kac denominator formula for the algebra $G = SL_2[z, z^{-1}]$ is

$$\prod_{n>0}(1 - q^{2n})(1 - q^{2n-1}z)(1 - q^{2n-1}z^{-1}) = \sum_{n\in\mathbf{Z}}(-1)^n q^{n^2} z^n.$$

This is the Jacobi triple product identity, and is also the Macdonald identity for the affine Lie root system corresponding to A_1.

Dyson described Macdonald's discovery of the Macdonald identities in [D]. Dyson found identities for $\eta(\tau)^m = q^{m/24}\prod_{n>0}(1 - q^n)^m$ for the following values of m:

$$3, 8, 10, 14, 15, 21, 24, 26, 28, \ldots$$

and wondered where this strange sequence of numbers came from. (The case $m = 3$ is just the Jacobi triple product identity with $z = 1$.) Macdonald found his identities corresponding to affine root systems, which gave an explanation for the sequence above: with one exception, the numbers are the dimensions of simple finite dimensional complex Lie algebras. The exception is the number 26 (found by Atkin), which as far as I know has not been explained in terms of Lie algebras. It seems possible that it is somehow related to the fake monster Lie algebra and the special dimension 26 in string theory.

Next we give a quick explanation of "completely replicable" functions. A function is called completely replicable if its coefficients satisfy certain relations. As an example of a completely replicable function, we will look at the elliptic modular function $j(\tau) - 744 = \sum c(n)q^n$. This satisfies the identity

$$j(\sigma) - j(\tau) = p^{-1}\prod_{\substack{m>0\\ n\in\mathbf{Z}}}(1 - p^m q^n)^{c(mn)}$$

where $p = e^{2\pi i\sigma}$, $q = e^{2\pi i\tau}$. (This formula was proved independently in the 80's by Koike, Norton, and Zagier, none of whom seem to have published their proofs.) Comparing coefficients of $p^m q^n$ on both sides gives many relations between the coefficients of j whenever we have a solution of $m_1 n_1 = m_2 n_2$ in positive integers, which are more or less the relations needed to show that j is completely replicable. For example, from the relation $2 \times 2 = 1 \times 4$ we get the relation

$$c(4) = c(3) + \frac{c(1)^2 - c(1)}{2}$$

or equivalently

$$20245856256 = 864299970 + \frac{196884^2 - 196884}{2}.$$

In the rest of this paper we will discuss various extensions of the original moonshine conjectures, some of which are still unproved. The first are Norton's "generalized moonshine" conjectures [N]. If we look at the Hauptmodul $T_{2A}(\tau) = q^{-1} + 4372q + \ldots$ we notice that one of the coefficients is almost the same as the dimension 4371 of the smallest nontrivial irreducible representation of the baby monster simple group, and the centralizer of an element of type 2A in the monster is a double cover of the baby monster. Similar things happen for other elements of the monster, suggesting that for each element g of the monster there should be some sort of graded moonshine module $V_g = \oplus_n V_{g,n}$ acted on by a central extension of the centralizer $Z_M(g)$. In particular we would get series $T_{g,h}(\tau) = \sum_n Tr(h|V_{g,n})q^n$ satisfying certain conditions. Some progress has been made on this by Dong, Li, and Mason [D-L-M], who proved the generalized moonshine conjectures in the case when g and h generate a cyclic group by reducing to the case when $g = 1$ (the ordinary moonshine conjectures). G. Höhn [H] has made some progress in the harder case when g and h do not generate a cyclic group by constructing the required modules for the baby monster (when g is of type $2A$). It seems likely that his methods would also work for the Fischer group Fi_{24}, but it is not clear how to go further than this. There might be some relation to elliptic cohomology (see [Hi] for more discussion of this), as this also involves pairs of commuting elements in a finite group and modular forms.

The space V_g mentioned above does not always have an invariant vertex algebra structure on it. Ryba discovered that a vertex algebra structure sometimes magically reappears when we reduce V_g modulo the prime p equal to the order of g. In fact V_g/pV_g can often be described as the Tate cohomology group $\hat{H}^0(g, V)$ for a suitable integral form V of the monster vertex algebra. This gives natural examples of vertex algebras over finite fields which do not lift naturally to characteristic 0. (Note that most books and papers on vertex algebras make the assumption that we work over a field of characteristic 0; this assumption is often unnecessary and excludes many interesting examples such as the one above.)

We will finish by describing some more of McKay's observations about the monster, which so far are completely unexplained. The monster has 9 conjugacy classes of elements that can be written as the product of two involutions of type $2A$, and their orders are 1, 2, 3, 4, 5, 6, 2, 3, 4. McKay pointed out that these are exactly the numbers appearing on an affine E_8 Dynkin diagram giving the linear relation between the simple roots. They are also the degrees of the irreducible representations of the binary icosahedral group. A similar thing happens for the baby monster: this time there are 5 classes of elements that are the product of two involutions of type $2A$ and their orders are 2, 4, 3, 2, 1. (This is connected with the fact that the baby monster is a "3,4-transposition group".) These are the numbers on an affine F_4 Dynkin diagram, and if we take the "double cover" of an F_4 Dynkin diagram we get an E_7 Dynkin diagram. The number on an E_7 Dynkin diagram are 1, 1, 2, 2, 3, 3, 4, 2 which are the dimensions of the irreducible representations of the binary octahedral group. The double cover of the baby monster is the centralizer of an element of order 2 in the monster. Finally a similar thing happens for $Fi_{24}.2$: this time there are 3 classes of elements that are the product of two involutions of type $2A$ and their orders are 2, 3, 1. (This is connected with the fact that $F_{24}.2$ is a "3-transposition group".) These are the numbers on an affine G_2 Dynkin diagram, and if we take the "triple cover" of an G_2 Dynkin diagram we get an E_6 Dynkin diagram. The number on an E_6 Dynkin diagram are 1, 1, 1, 2, 2, 2, 3, which are the dimensions of the irreducible representations of the binary tetrahedral group. The triple cover of $Fi_{24}.2$ is the centralizer of an element of order 3 in the monster.

The connection between Dynkin diagrams and 3-dimensional rotation groups is well understood (and is called the McKay correspondence), but there is no known explanation for the connection with the monster.

References.

[B94] R. E. Borcherds, Simple groups and string theory, First European congress of mathematics, Paris July 1992, Ed. A. Joseph and others,vol. 1 p. 411-421, Birkhauser 1994.

[B98] R. E. Borcherds, Automorphic forms and Lie algebras. Current developments in mathematics 1996, to be published by international press, 1998. Also available from www.dpmms.cam.ac.uk/~reb.

[C-G] C. J. Cummins, T. Gannon, Modular equations and the genus zero property of moonshine functions. Invent. Math. 129 (1997), no. 3, 413–443.

[C-N] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[D] F. J. Dyson, Missed opportunities. Bull. Amer. Math. Soc. 78 (1972), 635–652.

[D-L-M] C. Dong, H. Li, G. Mason, Modular invariance of trace functions in orbifold theory, preprint q-alg/9703016.

[F-L-M] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988. (Also see the announcement A natural representation of the Fischer-Griess monster with the modular function J as character, Proc. Natl. Acad. Sci. USA 81 (1984), 3256-3260.)

[Ge] R. W. Gebert, Introduction to vertex algebras, Borcherds algebras and the monster Lie algebra, Internat. J. Modern Physics A8 (1993) no 31, 5441-5503.

[Go] The Work of R.E. Borcherds, this volume, preprint math/9808136.

[G-T] P. Goddard and C. B. Thorn, Compatibility of the dual Pomeron with unitarity and the absence of ghosts in the dual resonance model, Phys. Lett., B 40, No. 2 (1972), 235-238.

[G] R. L. Griess, The friendly giant. Invent. Math. 69 (1982), no. 1, 1–102.

[Hi] F. Hirzebruch, T. Berger, R. Jung, Manifolds and modular forms. Aspects of Mathematics, E20. Friedr. Vieweg & Sohn, Braunschweig, 1992. ISBN: 3-528-06414-5

[H] G. Höhn, Selbstduale Vertexoperatorsuperalgebren und das Babymonster. [Self-dual vertex-operator superalgebras and the Baby Monster] Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, 1995. Bonner Mathematische Schriften 286. Universität Bonn, Mathematisches Institut, Bonn, 1996.

[J] E. Jurisich, Generalized Kac-Moody Lie algebras, free Lie algebras and the structure of the Monster Lie algebra. J. Pure Appl. Algebra 126 (1998), no. 1-3, 233–266.

[K90] V. G. Kac, "Infinite dimensional Lie algebras", third edition, Cambridge University Press, 1990. (The first and second editions (Birkhauser, Basel, 1983, and C.U.P., 1985) do not contain the material on generalized Kac-Moody algebras.)

[Ko] M. Koike, On Replication Formula and Hecke Operators, Nagoya University preprint.

[LZ] B. Lian, G. Zuckerman, New perspectives on the BRST-algebraic structure in string theory, hepth/921107, Communications in Mathematical Physics 154, (1993) 613-64, and Moonshine cohomology, q-alg/ 950101 Finite groups and Vertex Operator Algebras, RIMS publication (1995) 87-11.

[M] Y. Martin, On modular invariance of completely replicable functions. Moonshine, the Monster, and related topics (South Hadley, MA, 1994), 263–286, Contemp. Math., 193, Amer. Math. Soc., Providence, RI, 1996.

[N] S. P. Norton, Appendix to G. Mason's paper in The Arcata Conference on Representations of Finite Groups (Arcata, Calif., 1986), 181–210, Proc. Sympos. Pure Math., 47, Part 1, Amer. Math. Soc., Providence, RI, 1987.

Modular Forms and Moonshine
A Brief introduction

By David Padvorac

Abstract

Modular forms arise in primarily in algebra and number theory; as is well known, algebra and number theory are intricately interwoven. It is well known that algebra and number theory arise often in the natural world; likewise for modular forms, whose base is in both algebra and number theory. Elliptic curves arise in both algebra and number theory, and are tightly connected with modular forms. My goal will be to develop most of the necessary content related to modular forms to sketch the motivating observation behind the Monstrous Moonshine conjecture. The actual conjecture involves module theory and representation theory, which are both rather out of the scope of this project.

Introduction

As noted in the abstract, the end motivation here, in simple, is to develop the definitions needed to actually understand the very peculiar key observation that brought about “moonshine” theory, a recent development interweaving number theory, algebra, and to a lesser extent, physics. By taking this admittedly very shallow approach to introducing modular forms, it is sure that a large portion of their depth will be lost; as an undergraduate research project, this is nearly bound to happen regardless, as many needed ideas are in the toolboxes of graduate algebra. That is the downside of what I’ll be presenting here. The upshot is that it will be an undergraduate-accessible introduction to these topics, and a small glimpse into the motivation of the Monstrous Moonshine theory and conjectures.

Lattices

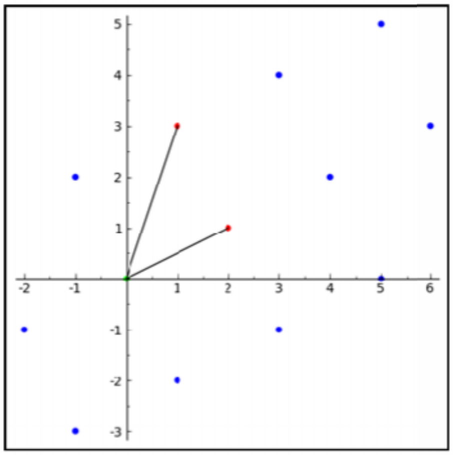
The lattice is a tremendously important tool; for starters, modular forms can be viewed as analytic functions on lattices in the complex plane. This will give us a simple definition to begin looking at modular forms with. A lattice λ in the complex plane is completely defined by two points in the plane with a non-real ratio:

$$\omega_1, \omega_2, \tau \in C; \frac{\omega_1}{\omega_2} = \tau \in H$$

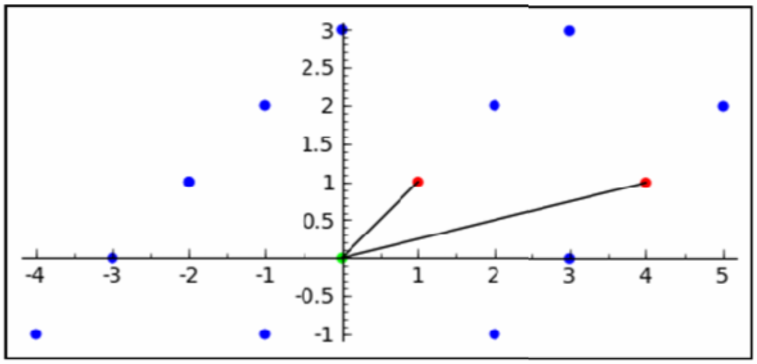
$$\lambda = L(\omega_1, \omega_2) = \{a\omega_1 + b\omega_2 : a, b \in Z\}$$

where H is the upper half plane, C is the complex plane, and in a shocking twist of events, we call the integers Z. This pair of points, the basis of the lattice, can be thought of as a pair of vectors in R^2. If the ratio of the two points is real, than equivalently the pair of vectors in R^2 is collinear, and will generate a boring lattice. Below are several examples of lattices, $L(1 + i, 4 + i)$ and $L(2 + i, 1 + 3i)$, plotted in Sage. The two generating points, ω_1 and ω_2 are red and the other lattice points are blue.

$L(2 + i, 1 + 3i)$



$L(1 + i, 4 + i)$



It is worth noting that if one of the basis points is in the lower half plane, it can be “shifted” to the upper half plane by an addition of an integer multiple of the other basis point, and doing this yields a basis that generates the same lattice. For simplicity’s sake, we’ll only use bases with both points in the upper half plane. Additionally, the lattice base does not have to be integer; integer bases were chosen for ease of visualization/calculation.

A function $G: \lambda \to C$ is simply a complex function restricted to the given set of lattice points; for example, taking the trivial lattice of Gaussian integers, $\lambda = L(1, i)$, $G(\lambda)$ is simply the range of G when evaluated at integer points in C. An important concept is the homogeneous function:

$G: \lambda \to C$ is homogeneous of degree –k, if:

$$G(a\lambda) = a^{-k}G(\lambda);\ \ k \in Z, a \in C$$

Examples:

1. $k = 0$: $G(a\lambda) = G(\lambda)$
2. $k = 1$: $G(a\lambda) = \frac{1}{a}G(\lambda)$
3. $k = 2$: $G(a\lambda) = \frac{1}{a^2}G(\lambda)$

It is worth noting that in these examples, that by intuition there seems to be a periodicity of some sort. Unfortunately a straight-up “normal” periodicity, as observed with sine or cosine, or any other similar function, only appears in a special case where $k = 0$; the other cases are this peculiar “scaling” sort of periodicity.

The set of homogenous functions of degree –k is in direct bijective correspondence with the set of modular forms of weight k; the proof is in Lang’s text, and will not be copied here. This simply gives an additional context for modular forms to be viewed in, which plays a key role later in the more detailed development of modular forms omitted in this paper.

The Modular Group

The modular group is a name for the group of invertible $n\ x\ n$ integer matrices, otherwise known as $SL_2(Z)$. Each element of the modular group is assigned to a function, a Linear Fractional Transformation, and operates on the upper half plane as follows:

$$\tau \in H\ ;\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \sigma \in SL_2(Z)$$
$$\sigma(\tau) = \frac{a\tau + b}{c\tau + d}$$

The proof that $\tau \in H \Rightarrow \sigma(\tau) \in H$, i.e. that the modular group maps the upper half plane to itself, is an elementary calculation; the only significant step is the clearing of complex numbers from the denominator of $\sigma(\tau)$, and this is done by multiplying the numerator and denominator by the complex conjugate of the denominator. The modular group is generated by two elements, $\sigma_1$andσ_2:

$$\sigma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix};\ \sigma_1(\tau) = \tau + 1$$
$$\sigma_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix};\ \sigma_2(\tau) = -\frac{1}{\tau}$$

The proof of this fact is a somewhat involved calculation but does not provide much deep insight, and so is omitted.

Some Complex Analysis Terminology

- A complex function is holomorphic if it is complex-differentiable (similar to differentiability of a function from R^2 to R^2) on some neighborhood of every point in its domain.
- Meromorphicity is a very slightly generalized version of holomorphicity. A function with isolated poles, which is holomorphic everywhere except at those poles, is called "meromorphic."

A nice intuitive geometry way to think of a meromorphic function is as the top of an old style camping tent; generally speaking, it's a nice and smooth surface except for where the tent poles poke the material up. Those points aren't too nice, but they are well isolated.

Weak Modularity

A meromorphic function $f: H \to C$ is considered weakly modular of weight k if the following condition holds:

$$f(\sigma(\tau))(c\tau + d)^k = f(\tau), \sigma(\tau) = \frac{a\tau+b}{c\tau+d}$$

where σ is an element of the modular group and τ is an element of the upper half plane. Because the modular group is generated by the afore mentioned two elements σ_1 and σ_2, we can check if any function is weakly modular of some weight by checking these two inequalities, which come from plugging σ_1 and σ_2 into our definition:

1. $\sigma_1(\tau) = \tau + 1$: $\quad f(\tau) = f(\tau + 1)$
2. $\sigma_2(\tau) = -\frac{1}{\tau}$: $\quad f(\tau) = \tau^k f\left(-\frac{1}{\tau}\right)$

There are several important facts we can observe.

1. For all weights:
    a. $f(\tau)$ is Z-periodic, i.e. it repeats itself on a unit interval: $f(\tau) = f(\tau + n)$ for any integer n. This can be proven by a trivial application of induction to (1).

   b. The zeroes and poles of $f(\tau)$ are invariant under operation by $SL_2(Z)$ on τ. This follows by comparing the right and left hand sides of the definition above: clearly the factor $(c\tau + d)^k$ will not have zeroes or poles, so $f(\tau)$ has zeroes and/or poles if corresponding exactly to those of $f\big(\sigma(\tau)\big)$.

2. For weight $k = 0$, $f(\tau)$ is invariant under the modular group, $SL_2(Z)$. This follows from setting $k = 0$ in equation (2).
3. For any odd weight $f(\tau) \equiv 0$. We can see this by setting $\sigma = -I$ in the definition. This gives us $f(\tau)(-1) = f(\tau)$, clearly only true when $f(\tau) \equiv 0$.
4. For weight $k = 2$, we get a result familiar from path integrals of complex analysis. There is a bit of legwork to do to see this result, however.
   a. First we manipulate our definition to get that $\frac{f(\sigma(\tau))}{f(\tau)} = \frac{1}{(c\tau+d)^2}$
   b. Next, we note that $\frac{d\sigma}{d\tau} = \frac{1}{(c\tau+d)^2}$; this derivative can be calculated easily using the quotient rule and the definition of $\sigma(\tau)$.
   c. Finally, equating the right hand sides of the two equations and redistributing the fractions, we have $f\big(\sigma(\tau)\big)d\sigma(\tau) = f(\tau)d\tau$, and so a path integral in the complex plane is invariant under $SL_2(Z)$.

Modular Forms

A function $f: H \to C$ is called a modular form of weight k if the following conditions satisfied:

1. f is holomorphic on H
2. f is holomorphic “at infinity”
3. f is weakly modular of weight k

For a function $f(\tau)$ to be holomorphic “at infinity” simply means that it must be bounded as $Im(\tau) \to \infty$. An easy way to show this would be by showing that $\lim_{Im(\tau)\to\infty} f(\tau)$ exists and is finite. This condition is necessary for some further results on the set of all modular forms of some given weight; this set is denoted $M_k(SL_2(Z))$.

An important example of specific type of modular form is the cusp form of weight k, a modular form of weight k, with the addition that in its Fourier expansion, the first terms coefficient is zero.

$$f(\tau) = \sum_{n=1}^{\infty} a_n q^n, \quad q = e^{2\pi i\tau}$$

$f(\tau)$ a cusp form □ $a_n = 0$

Also, from this definition we also have that a modular form is a cusp form when $\lim_{Im(\tau)\to\infty} f(\tau) = 0$. The significance of cusp forms to this work will be in defining a very important modular form in terms of them; this function plays a key role in the question that motivated the moonshine conjectures.

Eisenstein Series

An Eisenstein series is essentially a 2-dimensional analog of the Riemann zeta function. In good tradition (and it also happens to be the generally accepted standard definition), we first use a definition that looks just like Riemann's original definition (this definition and it's immediate corollaries are directly from Diamond and Shurman).

$$G_k(\tau) = \sum_{(c,d) \in I} \frac{1}{(c\tau + d)^k} ;$$

$$\tau \in \mathcal{H}$$
$$k \in \mathbb{Z}$$
$$I = \mathbb{Z}^2 - \{(0,0)\}$$

G_k is absolutely convergent $\forall k$ and uniformly convergent on compact subsets of $\mathcal{H}$. By definition than, G_k is holomorphic on $\mathcal{H}$. By applying an arbitrary element σ of $SL_2(Z)$ to τ and using some tricky manipulation of the summation of $G_k(\sigma(\tau))$, it falls out that G_k is weakly modular with weight k. Lastly, G_k is bounded as $Im(\tau) \to \infty$, so it is indeed a modular form.

We now define two important cusp forms. They do seem entirely arbitrary, and neither their purpose nor significance readily meets the eye. This is understandable, considering that they are part of a much more complicated thing, the Weierstrass $\wp$-function: they are the 2nd and 3rd coefficients of its Laurent expansion. The $\wp$-function won't be discussed here, although it is part of Koblitz's presentation of elliptic curves, and can be studied in full depth in his text.

$$g_2(\tau) = 60G_4(\tau)$$
$$g_3(\tau) = 140G_6(\tau)$$

Now, we define the elliptic discriminant function $\Delta(\tau)$, and the modular invariant $j(\tau)$. Again, like with g_2 and g_3, a full development of the nature and contexts of these functions isn't included here. They both play key roles in the theory of elliptic curves and modular forms.

$$\Delta: \mathcal{H} \to \mathbb{C}$$
$$\Delta(\tau) = (g_2(\tau))^3 + 27(g_3(\tau))^2$$

$$j: \mathcal{H} \to \mathbb{C}$$
$$j(\tau) = 1728 \frac{(g_2(\tau))^3}{\Delta(\tau)}$$

There is a trivially-proved result on modular forms that the weight of the product of two forms is the product of their weights. From this, we can note that the numerator and denominator of $j(\tau)$ both have the same weight; it follows that $j(\tau)$ is invariant under $SL_2(Z)$. Finally, writing $j(\tau)$ as a Fourier series in terms of $q = e^{2\pi i\tau}$:

$$j(\tau) = q^{-1} + 744 + 196{,}884q + 21{,}493{,}760q^2 + 864{,}299{,}970q^3 + \cdots$$

Note: the four definitions given here are taken straight from the corresponding section of the text by Diamond and Shurman.

This is the end of the good definitions I'll give here. The rest of the needed content is either from entry level algebra or upper level algebra, or just simply very inaccessible in any rigor to most typical undergraduates, myself included. For the inaccessible content, I'll give the hand-waving definitions and descriptions.

Some Light Definitions

1. Finite Simple Groups
   Gannon very succinctly describes finite simple groups in his text; I'll summarize it here:
   "Finite simple groups are to finite groups what the primes are to integers—they are their elementary building blocks."

   All finite simple groups have been classified, this monumental work was wrapped up, give or take, in the late 90's. There are 18 infinite families (for example, cyclic groups of prime order, $\mathbb{Z}/p\mathbb{Z}$) and 26 exceptional groups. The largest of these groups is the so called "monster" group of Fischer and Griess. The order of the monster group is approximately 8×10^{53}.

2. Representation theory
   Again, we'll use Gannon's description:
   A representation of a group is an assignment of an $n \times n$ matrix $A(g)$ to each element of that group, so that the matrix product respects the group product:
   $$A(g_1)A(g_2) = A(g_1g_2)$$
   In essence, a representation of a group is a full description of its structure in the language of linear algebra. The dimension of a representation is the size n of the matrices.
   An representation can be reduced in terms of subspaces of the vector space V acted on by the matrices A. A subspace of V fixed under the group action matrices A is essentially a building block for the representation.

The Observation!

And finally, we've come to the point where we can look at the observation made by John McKay, detailed in Gannon's text. Mckay first noticed the following:

$$196{,}884 \approx 196{,}883$$
$$21{,}493{,}760 \approx 21{,}296{,}876$$
$$864{,}299{,}970 \approx 842{,}609{,}326$$

and shortly came to a more precise resulting numerology:

$$196{,}884 = 196{,}883 + 1$$
$$21{,}493{,}760 = 21{,}296{,}876 + 196{,}883 + 1$$
$$864{,}299{,}970 = 842{,}609{,}326 + 21{,}296{,}876 + 2 \times 196{,}883 + 2 \times 1$$

The numbers on the right side of these equalities are the first coefficients of the modular invariant, $j(\tau)$. The numbers on the right hand side are the dimensions of smallest "building blocks" of the group representation of the Monster group.

This striking set of equalities continues, for all coefficients of the modular invariant. The conjecture this observation lead to is, again, far beyond the means of this paper. Regardless, I do believe that this seemingly-out-of-nowhere correspondence is simply another example of the relations that appear in seemingly completely unrelated fields of mathematics. There are implications of this observation in physics, more specifically string theory; even rudimentary explanations of these implications are far beyond my current grasp.

Bibliography

Diamond, F., & Shurman, J. (2005). A First Course in Modular Forms. Springer.

Gannon, T. (2006). Moonshine Beyond the Monster. Cambridge, UK: Cambridge University Press.

Koblitz, N. (1993). Introduction to Elliptic Curves and Modular Forms. Springer.

Lang, S. (1976). An Introduction to Modular Forms. New York City, NY: Springer.

Modular Moonshine II. *24 July 1994, corrected 19 Sept 1995*

Duke Math. J. 83 (1996) no. 2, 435-459.

Richard E. Borcherds,*
Mathematics department,
University of California at Berkeley,
CA 94720-3840
U. S. A.
e-mail: reb@math.berkeley.edu

Alex J. E. Ryba,
Dept of Mathematics,
Marquette University,
Milwaukee,
WI 53233,
U. S. A.
e-mail: alexr@sylow.mscs.mu.edu

The monster simple group acts on the monster vertex algebra, and the moonshine conjectures state that the traces of elements of the monster on the vertex algebra are Hauptmoduls. Ryba [R94] conjectured the existence of similar vertex algebras over fields of characteristics p acted on by the centralizers of certain elements of prime order p in the monster, and conjectured that the Brauer traces of p-regular elements of the centralizers were certain Hauptmoduls. We will prove these conjectures when the centralizer involves a sporadic group ($p \leq 11$, corresponding to the sporadic groups B, Fi'_{24}, Th, HN, He, and M_{12}).

Contents.

1. Introduction.
Notation.
2. Cohomology and modular representations.
3. Vertex superalgebras mod p.
4. A vanishing theorem for cohomology.
5. The case $p = 2$.
6. Example: the Held group.
7. Open problems and conjectures.

1. Introduction.

The original "moonshine conjectures" of Conway, Norton, McKay, and Thompson said that the monster simple group M has an infinite dimensional graded representation $V = \oplus_{n\in\mathbf{Z}} V_n$ such that the dimension of V_n is the coefficient of q^n of the elliptic modular function $j(\tau)-744 = q^{-1}+196884q+\cdots$, and more generally the McKay-Thompson series $T_g(\tau) = \sum_{n\in\mathbf{Z}} \mathrm{Tr}(g|V_n)q^n$ is a Hauptmodul for some genus 0 congruence subgroup of $SL_2(\mathbf{R})$. The representation V was constructed by Frenkel, Lepowsky, and Meurman [FLM], and it was shown to satisfy the moonshine conjectures in [B92] by using the fact that it carries the structure of a vertex algebra [B86], [FLM].

Meanwhile, Norton had suggested that there should be a graded space associated to every element g of the monster acted on by some central extension of the centralizer of g ([N], see also [Q]). It is easy to see that these graded spaces are usually unlikely to have a vertex algebra structure. Ryba suggested [R94] that these spaces might have a vertex algebra structure if they were reduced mod p (if g has prime order p). He also suggested the following definition for these vertex algebras:

$$^{g}V = \frac{V^g/pV^g}{(V^g/pV^g)\cap(V^g/pV^g)^{\perp}}$$

---

* Supported by NSF grant DMS-9401186.

where V is some integral form for the monster vertex algebra, and V^g is the set of vectors fixed by g. He conjectured that the dimensions of the homogeneous components of ${}^gV$ should be the coefficients of certain Hauptmoduls, and more generally that if h is a p-regular element of $C_M(g)$ then

$$\mathrm{Tr}(h|{}^gV_n) = \mathrm{Tr}(gh|V_n)$$

where the trace on the left is the Brauer trace. (The numbers on the right are known to be the coefficients of a Hauptmodul depending on gh by [B92].) We will call these conjectures the modular moonshine conjectures.

We now describe the proof of the modular moonshine conjectures for some elements g given in this paper. We observe (proposition 2.3) that the definition of ${}^gV$ in [R94] is equivalent to defining it to be the Tate cohomology group $\hat{H}^0(g, V)$, or at least it would be if a good integral form V of the monster Lie algebra was known to exist. The lack of a good integral form V does not really matter much, because we can just as well use a $\mathbf{Z}[1/n]$-form for any n coprime to $|g|$, and a $\mathbf{Z}[1/2]$-form can be extracted from Frenkel, Lepowsky and Meurman's construction of the monster vertex algebra.

It is difficult to work out the dimension of $\hat{H}^0(g, V[1/2]_n)$ directly. However it is easy to work out the difference of the dimensions of $\hat{H}^0$ and $\hat{H}^1$, which can be thought of as a sort of Euler characteristic (and is closely related to the Herbrand quotient in number theory). This suggests that we should really be looking at $\hat{H}^*(g, V[1/2]) = \hat{H}^0(g, V[1/2]) \oplus \hat{H}^1(g, V[1/2])$, and it turns out for essentially formal reasons that this has the structure of a vertex superalgebra. The traces of elements h of the centralizer of g on this superalgebra are just given by the traces of the elements gh on the monster vertex algebra.

We can now complete the proof of the modular moonshine conjectures for the element g of the monster by showing that $\hat{H}^1(g, V[1/2]) = 0$. We do this for certain elements of odd prime order of the monster coming from M_{24} in section 4; this is a long but straightforward calculation. There are plenty of elements g of the monster for which $\hat{H}^1(g, V[1/2])$ does not vanish; this happens whenever some coefficient of the Hauptmodul of g is negative; for example, g of type 3B. There are also some elements of large prime order for which we have been unable to prove the modular moonshine conjectures because we have not proved that $\hat{H}^1(g, V[1/2]) = 0$ (though this is probably true), but we do at least cover all the cases when g has prime order and its centralizer involves a sporadic simple group.

The case when $p = 2$ has several extra complications, due partly to the existence of several extensions of groups by 2-groups, and due partly to the fact the FLM construction can be carried out over $\mathbf{Z}[1/2]$ but it is not clear how to do it over $\mathbf{Z}$. We deal with these extra problems in section 5 by using a $\mathbf{Z}[1/3]$-form of the monster vertex algebra. To construct this $\mathbf{Z}[1/3]$-form we need to make a mild assumption (which we have not checked) about the construction of the monster vertex algebra from an element of order 3 announced by Dong and Mason and by Montague. (In particular the proof of the modular moonshine conjectures for elements of type 2A in the monster uses this assumption.)

In section 6 we give some calculations illustrating the case when g is an element of type 7A in the monster (so the centralizer $C_M(g)$ of g is $\langle g\rangle \times He$). We give the characters and the decomposition into irreducible modular representations of the first few graded pieces of $\hat{H}^0(g, V[1/2])$.

In section 7 we discuss some open problems, in particular how one might construct a good integral form of the monster vertex algebra.

Notation.

A^G	The largest submodule of A on which G acts trivially.
A_G	The largest quotient module of A on which G acts trivially.
A, B, C	G-modules.
Aut	The automorphism group of something.
B	The baby monster sporadic simple group.
$\mathbf{C}$	The complex numbers.
$C_M(g)$	The centralizer of g in the group M.
Fi_{24}	One of Fischer's groups.
$\mathbf{F_q}$	The finite field with q elements.
g	An element of G, usually of order p.
$\langle g\rangle$	The group generated by g.
G	A group, often cyclic of prime order p and generated by g.

$\hat{H}^i(G,A)$	A Tate cohomology group of the finite group G with coefficients in the G-module A.
$\hat{H}^i(g,A)$	means $\hat{H}^i(\langle g\rangle,A)$, where $\langle g\rangle$ is the cyclic group generated by g.
$\hat{H}^*(g,A)$	The sum of the Tate cohomology groups $\hat{H}^0(g,A)$ and $\hat{H}^1(g,A)$, considered as a super module.
He	The Held sporadic simple group.
HN	The Harada-Norton sporadic simple group.
I	The elements of a group ring of norm 0.
Im	The image of a map.
Ker	The kernel of a map.
$\Lambda, \hat{\Lambda}$	The Leech lattice and a double cover of the Leech lattice.
L	An even lattice.
M	The monster simple group.
M_{12}	A Mathieu group.
M_{24}	A Mathieu group.
N	The norm map: $N(v)=\sum_{g\in G} g(v)$.
$N_M(g)$	The normalizer of the subgroup $\langle g\rangle$ in the group M.
p	A prime, usually the order of g.
$\mathbf{R}$	The real numbers.
R_p	A finite extension of the p-adic integers.
$S(a)$	$=\sum_{0\le n<p} ng^n(a)$
SL_2	A special linear group.
Th	Thompson's sporadic simple group.
Tr	$\mathrm{Tr}(g\vert A)$ is the usual trace of g on a module A if A is a module over a ring of characteristic 0, and the Brauer trace if A is a module over a field of finite characteristic.
$V[1/n]$	A $\mathbf{Z}[1/n]$-form of the monster vertex algebra.
V_Λ	The integral form of the vertex algebra of $\hat{\Lambda}$.
V_n	The degree n piece of V.
V^n	An eigenspace of some group acting on V.
${}^gV$	A modular vertex algebra or superalgebra given by $\hat{H}^*(g,V[1/n])$ for some n coprime to $\vert g\vert$.
$\mathbf{Z}$	The integers.
$\mathbf{Z}_p$	The p-adic integers.
ω	A cube root of 1 or a conformal vector.

2. Cohomology and modular representations.

In this section we summarize some basic facts about group cohomology that we will use later. For more details about Tate cohomology groups see the article [AW].

If G is a finite group acting on an abelian group A we define A^G to be the elements of A fixed by G, and A_G to be the largest quotient of A on which G acts trivially. We write N for the norm map defined by $N(a)=\sum_{g\in G} g(a)$. If n is any integer then $\hat{H}^n(G,A)$ is the Tate cohomology group of G with coefficients in A. It has the following properties.

1. $\hat{H}^0(G,A)=A^G/\mathrm{Im}(N)$.
2. $\hat{H}^{-1}(G,A)=\mathrm{Ker}(N|A_G)$.
3. If $0\to A\to B\to C\to 0$ is an exact sequence of G modules then
$$\cdots\longrightarrow \hat{H}^{i-1}(G,C)\longrightarrow \hat{H}^i(G,A)\longrightarrow \hat{H}^i(G,B)\longrightarrow \hat{H}^i(G,C)\longrightarrow \hat{H}^{i+1}(G,A)\longrightarrow \hat{H}^{i+1}(G,B)\longrightarrow\cdots$$
is exact.
4. There is a bilinear cup product from $\hat{H}^i(G,A)\times\hat{H}^j(G,B)$ to $\hat{H}^{i+j}(G,C)$, defined whenever we have a G-invariant bilinear map from $A\times B$ to C, and which is the obvious induced product when $i=j=0$. $\hat{H}^*(G,\mathbf{Z})$ is a super-commutative ring under cup product. More generally if A has any G-invariant algebraic structure defined in terms of multilinear products and multilinear identities (e.g., commutative ring, Lie algebra, vertex algebra, etc.) then $\hat{H}^*(G,A)=\oplus_{n\in\mathbf{Z}}\hat{H}^n(G,A)$ is a "super" version of this algebraic structure under the cup product; this means that it satisfies the same identities as A except that we insert a factor of -1 whenever we interchange the order of two elements of odd degree. This follows from the associativity and supercommutativity properties of the cup product given in [AW]. For

example, if A is a Lie algebra, then $\hat{H}^*(G,A)$ is a Lie superalgebra (at least if A is 2-divisible so that the Lie algebra axioms can all be written as multilinear identities).

5. If G is a cyclic group generated by an element g we will also write the cohomology groups $\hat{H}^i(G,A)$ as $\hat{H}^i(g,A)$. In this case $\hat{H}^2(g,\mathbf{Z})$ has a canonical element (depending on the generator g) such that cup product by this element is an isomorphism from $\hat{H}^i(g,A)$ to $\hat{H}^{i+2}(g,A)$. In particular there are essentially only 2 different cohomology groups, $\hat{H}^0$ and $\hat{H}^1$, as all the others are isomorphic to one of these two, so we define $\hat{H}^*(g,A)$ to be $\hat{H}^0(g,A)\oplus\hat{H}^1(g,A)$. As in 4 above, if A has some algebraic structure, then $\hat{H}^*(g,A)=\hat{H}^0(g,A)\oplus\hat{H}^1(g,A)$ has a super version of this algebraic structure, with $\hat{H}^0$ and $\hat{H}^1$ being the even and odd parts. Notice that the algebra product from $\hat{H}^1\times\hat{H}^1$ to $\hat{H}^0$ depends on the choice of generator g because it uses the isomorphism from $\hat{H}^2$ to $\hat{H}^0$. In particular if $g\in G$ and A has an algebra structure invariant under G, then the natural action of the normalizer $N_G(g)$ on $\hat{H}^*(g,A)$ need not preserve the algebra structure (unless of course $\hat{H}^1(g,A)=0$). The centralizer $C_G(g)$ always preserves the algebra structure. We always regard $\hat{H}^*(g,A)$ as a supermodule, and in particular when we take the trace of an element h on it we multiply the trace on $\hat{H}^1$ by -1, i.e., $\mathrm{Tr}(h|\hat{H}^*(g,A))$ is defined to be $\mathrm{Tr}(g|\hat{H}^0(g,A))-\mathrm{Tr}(g|\hat{H}^1(g,A))$. (The trace will always be the Brauer trace, which takes values in a field of characteristic 0.)
6. $\hat{H}^*(G,A)$ is a $|G|$-torsion module, and in particular if multiplication by $|G|$ is an isomorphism on A then $\hat{H}^*(G,A)$ vanishes.

We let $\mathbf{Z}_p$ be the ring of p-adic integers.

Lemma 2.1. *If A is a free $\mathbf{Z}[1/n]$ module acted on by a p-group G with $(p,n)=1$ then the natural map from $\hat{H}^*(G,A)$ to $\hat{H}^*(G,A\otimes\mathbf{Z}_p)$ is an isomorphism.*

Proof. This follows by looking at the long exact sequence of $0\to A\to A\otimes\mathbf{Z}_p\to A\otimes(\mathbf{Z}_p/\mathbf{Z}[1/n])\to 0$, if we observe that multiplication by p is an isomorphism on $\mathbf{Z}_p/\mathbf{Z}[1/n]$ so the cohomology of $A\otimes(\mathbf{Z}_p/\mathbf{Z}[1/n])$ vanishes. This proves lemma 2.1.

Now suppose that G is a cyclic group of order p generated by g. Recall ([CR] p. 690) that there are exactly 3 indecomposable finitely generated modules over the group ring $\mathbf{Z}_p[G]$ which are free as $\mathbf{Z}_p$ modules; these are:

1. The 1-dimensional module $\mathbf{Z}_p$ (with g acting trivially). The cohomology groups of G with coefficients in $\mathbf{Z}_p$ are $\hat{H}^0(g,\mathbf{Z}_p)=\mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g,\mathbf{Z}_p)=0$, as can be easily checked by explicit calculation.
2. The group ring $\mathbf{Z}_p[G]$. All cohomology groups of this module are trivial as it is a projective module.
3. The module I, which is the kernel of the natural map from $\mathbf{Z}_p[G]$ to $\mathbf{Z}_p$. The cohomology groups of I can be worked out from the long exact sequence of $0\to I\to\mathbf{Z}_p[G]\to\mathbf{Z}_p\to 0$, and are given by $\hat{H}^0(g,I)=0$, $\hat{H}^1(g,I)=\mathbf{Z}/p\mathbf{Z}$.

The next proposition is the main tool for calculating the Brauer characters of the modular vertex algebras we will construct.

Proposition 2.2. *Suppose that A is a finitely generated free module over $\mathbf{Z}_p$ (or over $\mathbf{Z}$) acted on by a group G containing an element g of order p in its center. Then the Brauer character of the virtual modular representation $\hat{H}^*(g,A)=\hat{H}^0(g,A)-\hat{H}^1(g,A)$ of $G/\langle g\rangle$ is given by*

$$\mathrm{Tr}(h|\hat{H}^*(g,A))=\mathrm{Tr}(gh|A)$$

for any p-regular element h of G. (The left hand trace is the Brauer trace, and the right hand one is the usual trace.)

Proof. We can assume that G is the product of the cyclic groups generated by g and h. We adjoin the $|h|$'th roots of 1 to $\mathbf{Z}_p$ to obtain an unramified extension R_p of $\mathbf{Z}_p$. It is sufficient to prove the proposition for the module $A\otimes R_p$ because both sides are unaffected by tensoring A by R_p. (The Brauer trace of the cohomology groups is of course calculated by regarding them as modules over the residue class field of R_p.) We can decompose $A\otimes R_p$ into eigenspaces for h, and as both traces are additive on modules we can assume that h acts on $A\otimes R_p$ as multiplication by some fixed root of unity ζ.

Both sides of proposition 2.2 are additive on short exact sequences (the additivity for the left hand side follows from the "exact hexagon" of Tate cohomology). The only irreducible representations of $Q\otimes R_p[g]$ over the quotient field $Q\otimes R_p$ of R_p are $Q\otimes R_p$ and $Q\otimes I\otimes R_p$, which implies that the module A can be built from $I \otimes R_p$, R_p, and finite modules by taking repeated extensions. By additivity it is then sufficient to prove proposition 2.2 in the cases when A is $I \otimes R_p$ or R_p or a finite module. The case when A is finite is trivial as both sides are zero (the dimensions of $\hat{H}^0(g, A)$ and $\hat{H}^1(g, A)$ are equal because the Herbrand quotient $|\hat{H}^0(g, A)|/|\hat{H}^0(g, A)|$ is 1 for finite modules A).

We have reduced the proof of proposition 2.2 to checking the 2 cases when $A \otimes R_p$ is R_p or $I \otimes R_p$, and where h acts as multiplication by ζ. If $A \otimes R_p$ is R_p, then $\hat{H}^0(g, R_p) = R_p/pR_p$, and $\hat{H}^1(g, A \otimes R_p) = 0$, so $\mathrm{Tr}(h|\hat{H}^*(R_p)) = \zeta$, which is equal to $\mathrm{Tr}(gh|R_p) = \mathrm{Tr}(h|R_p)$. If $A \otimes R_p$ is the group ring $R_p[g]$, then both cohomology groups are 0 and the trace of gh on $R_p[g]$ is also 0 as the eigenvalues of gh are ζ multiplied by every p'th root of 1 (including 1). Finally the case when $A \otimes R_p = I \otimes R_p$ follows from the previous two cases and the exact sequence

$$0 \longrightarrow I \otimes R_p \longrightarrow R_p[g] \longrightarrow R_p \longrightarrow 0$$

because both traces are additive on short exact sequences. This proves proposition 2.2.

If A is a finitely generated free module over a ring R with a bilinear form $(,)$, then we say this form is self dual if every linear map f from A to R is of the form $f(a) = (a, b)$ for some fixed $b \in A$. If the form is symmetric and R is the ring of integers this is equivalent to saying that A is a unimodular lattice. We say the bilinear form is nonsingular if there is no nonzero element a with $(a, b) = 0$ for all $b \in A$. We say that a submodule B of A is primitive if A/B is torsion free.

Proposition 2.3. *Suppose A is a free $\mathbf{Z}$ or $\mathbf{Z}_p$ module acted on by the finite group G of order p with an invariant self dual symmetric bilinear form $(,)$. Then $N(A)$ is the set of all elements $a \in A^G$ such that $p|(a, b)$ for all $b \in A^G$. In particular the bilinear form $(,)$ induces a nonsingular bilinear form on $\hat{H}^0(G, A)$.*

Proof. We will prove this when A is a free $\mathbf{Z}$-module; the proof when A is a free $\mathbf{Z}_p$ module is similar. We think of A as a lattice in the real vector space $A \otimes \mathbf{R}$, which also has a nonsingular bilinear form induced by that of A. Note that $A \otimes \mathbf{R}$ is the orthogonal direct sum of the subspaces $A^G \otimes \mathbf{R}$ and $\mathrm{Ker}(N)$, so the bilinear form is nonsingular on each of these subspaces. The map N is just p times orthogonal projection onto the subspace $(A \otimes \mathbf{R})^G$. Hence $\frac{1}{p}N(A)$ is just the orthogonal projection of A into $(A \otimes \mathbf{R})^G$. The sublattice A^G is also primitive because if na is fixed by G for some nonzero $n \in \mathbf{Z}$ then so is a. On the other hand, the set of vectors $a \in A^G$ such that $p|(a, b)$ for all $b \in A^G$ is the same as $p(A^{G'})$ (where B' means the dual of the lattice B, i.e., the set of all vectors of $B \otimes \mathbf{R}$ which have integral inner product with all elements of B). For any nonsingular primitive sublattice B of A, and in particular for $B = A^G$, B' is the projection of A into $B \otimes \mathbf{R}$ because the bilinear form on A is self dual. (As B is a primitive sublattice of A the natural map from A' to B' is surjective, so as A is self dual and can be identified with A' we see that the projection from A to B' is surjective.) This proves proposition 2.3.

There is a similar but slightly more complicated result for $\hat{H}^1$.

Proposition 2.4. *Suppose A is a free $\mathbf{Z}$ or $\mathbf{Z}_p$ module with an invariant self dual symmetric bilinear form $(,)$ acted on by the finite group G generated by g of order p. Define a bilinear form $\langle a, b\rangle$ on A by $\langle a, b\rangle = (S(a), b)$ where $S(a) = \sum_{0 \le n < p} ng^n(a)$. Then A_G is the quotient of A by the set C of all elements $a \in \mathrm{Ker}(N)$ such that $p|\langle a, b\rangle$ for all $b \in \mathrm{Ker}(N)$. In particular the bilinear form $\langle,\rangle$ induces a nonsingular antisymmetric bilinear form on $\hat{H}^1(G, A)$.*

Proof. Some easy calculations show (even if we do not assume $(,)$ is self dual) that $\langle,\rangle$ is antisymmetric mod$p$, that $S(a) - S(g(a)) = N(a) - pa$, that $N(S(a)) = (p-1)pN(a)/2$, and that the kernel of the map from A to A_G (i.e., the module generated by elements of the form $a - g(a)$) is contained in C.

Suppose $a \in C$; we want to deduce that $a = b - g(b)$ for some $b \in A$. We know that $N(S(a)) = 0$ (because $N(a) = 0$), and $p|(S(a), b)$ for all b with $N(b) = 0$. Hence $S(a)/p$ lies in $\mathrm{Ker}(N)'$. The submodule $\mathrm{Ker}(N)$ is a primitive nonsingular submodule of the self dual module A, so $S(a)/p$ lies in the orthogonal projection of A into $\mathrm{Ker}(N) \otimes \mathbf{R}$, i.e., $S(a)/p = b - N(b)/p$ for some $b \in A$. But this means that $S(a) = S(g(b)) - S(b)$, so that $S(a - b + g(b)) = 0$. But S is injective, as we can see by checking that it multiplies every eigenvector of g by a nonzero constant, so $a - b + g(b) = 0$. This proves proposition 2.4. (It is easy to see that $\hat{H}^1(G, A)$ has an antisymmetric $\mathbf{Z}/p\mathbf{Z}$-valued bilinear form induced from the cup product and the bilinear form on A, followed by the isomorphism from $\hat{H}^2(G, \mathbf{Z})$ to $\hat{H}^0(G, \mathbf{Z})$. The main point of the proof above is that this antisymmetric form is self dual when the form on A is self dual.)

3. Vertex superalgebras mod p.

In this section we will construct a vertex superalgebra for every element of odd prime order in the monster.

If R is a subring of the complex numbers $\mathbf{C}$, then by an R-form of the monster vertex algebra we mean a graded vertex algebra defined over R which is a free R module, and which becomes isomorphic to the FLM monster vertex algebra when it is tensored with $\mathbf{C}$, and which is acted on by the monster and contains a conformal vector (the components of whose vertex operator generate the Virasoro algebra). We say that an R-form is self dual if the natural bilinear form on the FLM algebra restricts to a self dual R-valued bilinear form on each homogeneous piece.

We let $V[1/2]$ be the $\mathbf{Z}[1/2]$-form of the monster vertex algebra constructed in theorem 3.2 below.

Theorem 3.1. *If g is an element of the monster of odd prime order p then $\hat{H}^*(g, V[1/2]) = \hat{H}^0(g, V[1/2]) \oplus \hat{H}^1(g, V[1/2])$ is a $\mathbf{Z}$-graded vertex superalgebra over $\mathbf{Z}/p\mathbf{Z}$, acted on by $C_M(g)/\langle g\rangle$. It has a nonsingular supersymmetric bilinear form that is invariant under $C_M(g)/\langle g\rangle$. If $h \in C_M(g)$ is p-regular then $\sum_{n\in\mathbf{Z}} \mathrm{Tr}(h|\hat{H}^*(g, V[1/2]_n))q^n$ is a Hauptmodul for some genus 0 subgroup of $SL_2(\mathbf{R})$, and is equal to the Hauptmodul of the element gh of M.*

Proof. The fact that $\hat{H}^*(g, V[1/2])$ is a vertex superalgebra follows from the remarks about group cohomology in the previous section. The nonsingularity of the supersymmetric bilinear form follows from propositions 2.3 and 2.4. The fact that the trace of g is a Hauptmodul follows from proposition 2.2 together with the result of [B92] that the series $\sum_{n\in\mathbf{Z}} \mathrm{Tr}(g|V_n)q^n$ is a Hauptmodul for any element $g \in M$. This proves theorem 3.1.

Remark. The bilinear form in theorem 3.1 is also compatible with the vertex algebra structure (where this notion is defined in the discussion below), but we will not use this fact.

Proposition 2.3 also shows that $\hat{H}^0(g, V[1/2])$ is isomorphic to the space ${}^gV = V^g/pV^g/((V^g/pV^g) \cap (V^g/pV^g)^\perp)$ suggested in [R94].

In the rest of this section we recall some results about vertex algebras from [B86] and give the construction of the $\mathbf{Z}[1/2]$-form $V[1/2]$ of the monster vertex algebra. The only serious use we make of vertex algebra structures is to use them to construct other modular vertex algebras, so the reader who wishes to ignore vertex algebras completely and only see the proof that the traces of various group elements are Hauptmoduls can replace the words "vertex algebra" by "graded vector space" without losing much. Similarly the conformal vectors and bilinear forms on vertex algebras are used mainly to construct conformal vectors and bilinear forms on modular vertex algebras, so the reader can ignore all the discussion of these.

We recall the definition of a vertex algebra over any commutative ring R. The definition we give here is equivalent to that given in [B86]. If V is an R module we define a vertex operator $a(z)$ on V to be a sequence of linear operators a_n for $n \in \mathbf{Z}$ on V, with the property that for any $b \in V$, $a_nb = 0$ for n sufficiently large (depending on a and b). We put these operators into the formal Laurent series $a(z) = \sum_{n\in\mathbf{Z}} z^n a_{-n-1}$. Then a vertex algebra over R is an R module V together with an element $1 \in V$ and a vertex operator $a(z)$, acting on V, for each $a \in V$, such that the following conditions are satisfied for any $a, b \in V$.

1. $1(z) = 1$, and $a(0)1 = a$ (i.e., $a_n1 = 0$ if $n \geq 0$ and $a_{-1}1 = a$).
2. The vertex operators $a(z)$ and $b(y)$ formally commute, in the sense that we can find $n \geq 0$ (depending on a and b) such that all coefficients of monomials x^iy^j of

$$(x - y)^n(a(x)b(y) - b(y)a(x))$$

   are zero. (For a vertex superalgebra we replace $b(y)a(x)$ by $(-1)^{(\deg(a),\deg(b))}b(y)a(x)$.)
3. We denote the element $a_{-n-1}1$ by $D^{(n)}(a)$. Then $(D^{(n)}a)(z) = ((\frac{d}{dz})^n/n!)a(z)$.

This is similar to a definition of commutative associative algebras over R: we can define these as R modules V such that for every $a \in V$ we are given an operator a_L on V and an element $1 \in V$ with the properties that $1_La = a_L1 = a$, and the operators a_L and b_L commute for all $a, b \in V$. This shows that a vertex algebra can be thought of as a "commutative ring with singularities in the multiplication".

The operators $D^{(n)}$ have the properties $D^{(0)}(a) = a$, $D^{(m)} = 0$ if $m < 0$, and $D^{(m)}D^{(n)} = \binom{m+n}{m}D^{(m+n)}$.

We briefly recall the structure of the integral form V_Λ of the vertex algebra of Λ from [B86]. The lattice Λ has a certain nonsplit central extension $\hat{\Lambda}$ (in which the group operation is written multiplicatively) such that if we denote a lifting of $a \in \Lambda$ by $e^a \in \hat{\Lambda}$, then $e^a e^b = (-1)^{(a,b)} e^b e^a$ so we get an exact sequence

$$0 \longrightarrow \pm 1 \longrightarrow \hat{\Lambda} \longrightarrow \Lambda \longrightarrow 0.$$

The automorphism group of $\hat{\Lambda}$ preserving the inner product on Λ is a nonsplit extension of the form $2^{24}.\mathrm{Aut}(\Lambda)$ which acts on the vertex algebra V_Λ of Λ (which really depends on $\hat{\Lambda}$ rather than Λ). V_Λ is graded by Λ in such a way that e^a has degree $a \in \Lambda$, it has an underlying ring structure which is "graded commutative", i.e., $ab = (-1)^{(\deg(a),\deg(b))} ba$ if a and b are homogeneous elements with Λ-degrees $\deg(a)$ and $\deg(b)$, and it has a derivation with divided powers, i.e., a set of additive operations $D^{(m)}$ for $m \in \mathbf{Z}$ such that $D^{(0)}(a) = a$, $D^{(m)}D^{(n)} = \binom{m+n}{m} D^{(m+n)}$, $D^{(m)} = 0$ if $m < 0$, and $D^{(n)}(ab) = \sum_{m \in \mathbf{Z}} (D^{(m)}(a)) D^{(n-m)}(b)$. The derivation D has Λ-degree 0; in other words $D^{(m)}(a)$ has the same Λ-degree as a. The graded-commutative ring V_Λ contains a copy of the twisted group ring of Λ which has a basis of elements denoted by e^a for $a \in \Lambda$ with the properties that $e^a e^b = \pm e^{a+b}$, $e^a e^b = (-1)^{(a,b)} e^b e^a$, and can be defined as the universal graded-commutative ring with a derivation with divided powers generated by this group ring. The subring $V_{\Lambda,0}$ of elements of V_Λ with Λ-degree 0 is generated as a commutative ring by elements of the form $(e^a)^{-1} D^{(i)}(e^a)$ for $a \in \Lambda$, $i \geq 1$.

A conformal vector of dimension (or "central charge") $c \in \mathbf{R}$ of a vertex algebra V is defined to be an element ω of V such that $\omega_0 v = D(v)$ for any $v \in V$, $\omega_1 \omega = 2\omega$, $\omega_3 \omega = c/2$, $\omega_i \omega = 0$ if $i = 2$ or $i > 3$, and any element of V is a sum of eigenvectors of the operator $L_0 = \omega_1$ with integral eigenvalues. If v is an eigenvector of L_0, then its eigenvalue is called the (conformal) weight of v. If v is an element of the monster vertex algebra V of conformal weight n, we say that v has degree $n-1$ $(= n - c/24)$, and we write V_n for the module of elements of V of degree n.

The $\mathbf{R}$-form of the vertex algebra of any c-dimensional even lattice L has a canonical conformal vector $\omega = \sum_i a_i(1)a_i(1)/2$ of dimension c, where the elements a_i run over an orthonormal basis of $L \otimes \mathbf{R}$ and the monster vertex algebra has a conformal vector of dimension 24. If ω is a conformal vector of a vertex algebra V then we define operators L_i on V for $i \in \mathbf{Z}$ by

$$L_i = \omega_{i+1}.$$

These operators satisfy the relations

$$[L_i, L_j] = (i-j)L_{i+j} + \binom{i+1}{3} \frac{c}{2} \delta^i_{-j}$$

and so make V into a module over the Virasoro algebra. The operator L_{-1} is equal to D.

The vertex algebra of any even lattice L has a real valued symmetric bilinear form (,) such that the adjoint of the operator u_n is $(-1)^i \sum_{j \geq 0} L_1^j (\sigma(u))_{2i-j-n-2}/j!$ if u has degree i, where σ is the automorphism of the vertex algebra defined by $\sigma(e^w) = (-1)^{(w,w)/2}(e^w)^{-1}$ for e^w an element of the twisted group ring of L corresponding to the vector $w \in L$. Similarly the monster vertex algebra has a real valued symmetric bilinear form (,) such that the adjoint of the operator u_n is $(-1)^i \sum_{j \geq 0} L_1^j (u)_{2i-j-n-2}/j!$ if u has degree i. If a vertex algebra has a bilinear form with the properties above we say that the bilinear form is compatible with the conformal vector. If a vertex algebra does not have a conformal vector but only a $\mathbf{Z}$-grading we can still define compatible bilinear forms, because we can define the operator $L_1^j/j!$ to be the adjoint of the operator $L_{-1}^i/i! = D^{(i)}$. (If we are not in characteristic 0 we have to modify these definitions slightly by replacing $L_1^i/i!$ by a system of divided powers of L_1 satisfying some conditions.)

The integral form V_Λ of the vertex algebra of the Leech lattice contains the conformal vector $\omega = \sum_{1 \leq i \leq 24} a_i(1)a_i(1)/2$ where the a_i's run over an orthogonal basis of $\Lambda \otimes \mathbf{R}$ because we can rewrite this as $\omega = \sum_{1 \leq i < j \leq 24} a_i(1)a_j(1)(a'_i, a'_j) + \sum_{1 \leq i \leq 24} a_i(1)a_i(1)(a'_i, a'_i)/2$ where the a_i's run over a basis of Λ and the a'_i's are the dual basis. The elements a'_i have integral inner products and even norms because they all lie in Λ as Λ is unimodular. In general the same argument shows that the conformal vector of any even integral unimodular lattice lies in the integral form of its vertex algebra; this is usually false for lattices that are not unimodular and even. The bilinear form on the integral form V_Λ is self dual (section 2, (vi) of [B86]), and in fact the same is true if Λ is replaced by any even unimodular lattice.

For working out values of the bilinear form in V_Λ it is useful to note that $(a(1)b(1), c(1)d(1)) = (a,c)(b,d) + (a,d)(b,c)$ for $a, b, c, d \in \Lambda$. In particular the conformal vector ω has norm $(\omega, \omega) = \dim(\Lambda)/2 = 12$.

Theorem 3.2. *There is a* $\mathbf{Z}[1/2]$*-form* $V[1/2]$ *of the monster vertex algebra with a conformal vector and a self dual bilinear form such that* $V[1/2]$ *can be written as a direct sum of* $2^{12}.M_{24}$ *modules each of which is isomorphic to a submodule of the* $\mathbf{Z}[1/2]$*-form* $V_\Lambda[1/2]$ *of* V_Λ*.*

Proof. Frenkel, Lepowsky and Meurman construct the monster vertex algebra as a complex vector space. However their construction almost works over the integers (if we use the integral form of V_Λ above) except that they often need to decompose vector spaces into eigenspaces of abelian groups of order 2. This can be done over $\mathbf{Z}[1/2]$ (proof: $a = \frac{a+h(a)}{2} + \frac{a-h(a)}{2}$). A precise definition of this $\mathbf{Z}[1/2]$-form $V[1/2]$ of the monster vertex algebra is given implicitly in the next 2 paragraphs.

The $\mathbf{Z}[1/2]$-form $V[1/2]$ of the monster vertex algebra splits as the sum $V^0 \oplus V^1$ of the $+1$ and -1 eigenspaces of an element $h \in M$ of type 2B. As modules over $\mathrm{Aut}(\hat{\Lambda}) = 2^{24}.\mathrm{Aut}(\Lambda)$, V^0_Λ and V^0 are isomorphic, where V^0_Λ is the subalgebra of $V_\Lambda[1/2]$ fixed by the automorphism σ.

To construct the action of the monster on $V[1/2]$, Frenkel, Lepowsky, and Meurman split it further as $V[1/2] = V^{00} \oplus V^{01} \oplus V^{10} \oplus V^{11}$ where the V^{ij}'s are the eigenspaces of a 2^2-group in the monster containing h, and can be taken as free $\mathbf{Z}[1/2]$ modules. The modules V^0 and V^1 decompose as $V^0 = V^{00} \oplus V^{01}$, $V^1 = V^{10} \oplus V^{11}$. Frenkel, Lepowsky, and Meurman construct a triality automorphism σ that maps V^{00} to itself, transitively permutes the modules V^{01}, V^{10}, and V^{11}, and commutes with the elements of the group $2^{12}.M_{24} \subset \mathrm{Aut}(\hat{\Lambda})$. We use the isomorphism defined by σ between these 3 modules to transport the $\mathbf{Z}[1/2]$-form from V^{01} to V^{10}, and V^{11}, which defines the $\mathbf{Z}[1/2]$-form $V[1/2] = V^{00} \oplus V^{01} \oplus V^{10} \oplus V^{11}$. In particular σ commutes with $2^{12}.M_{24}$, so that V^{01}, V^{10}, and V^{11} are all isomorphic as modules over this group.

The conformal vector of $V[1/2]$ is just the conformal vector of $V^0 \subset V_\Lambda[1/2]$ (which is fixed by h). The self dual bilinear form on V_Λ induces a self dual bilinear form on V^0 over $\mathbf{Z}[1/2]$, which restricts to self dual forms on V^{01} and V^{00}, which give self dual forms on V^{10} and V^{11} by using the triality automorphism to transport the self dual form from V^{01}. This defines the self dual form on $V[1/2]$. This proves theorem 3.2.

4. A vanishing theorem for cohomology.

For every element g of the monster (or at least for those of odd order) we have constructed a vertex superalgebra $\hat{H}^*(g, V)$ acted on by $C_M(g)$ whose McKay-Thompson series are Hauptmoduls. For some elements g we will now show that the first cohomology group vanishes, so in fact we have constructed a modular vertex algebra satisfying the modular moonshine conjectures for the element g. It is certainly not true that $\hat{H}^1(g, V)$ vanishes for all elements g; by proposition 2.2 a necessary condition is that all coefficients of the Hauptmodul of g are nonnegative. This may be a sufficient condition but we do not know how to prove this in general. We will prove it is sufficient if g is an element of odd order coming from an element of M_{24}. This covers the cases of the conjugacy classes 3A, 3C, 5A, 7A, 11A, 15A, 21A, and 23A of the monster. In section 5 we will prove a similar theorem for elements of type 2A (under a mild assumption).

We will call a free R-module A acted on by a group G a permutation module if it has a basis that is closed under the action of G.

Lemma 4.1. *If* A *is a permutation module of a finite group* G *over a ring* R *with no* $|G|$ *torsion then* $\hat{H}^{-1}(G, A) = 0$*.*

Proof. We can assume that A comes from a transitive permutation representation of G, as all permutation representations are sums of transitive ones. Suppose S is the set of elements in the permutation representation, so that S is a basis for A. We have to show that $\mathrm{Ker}(N)$ is spanned by elements of the form $a - g(a)$. But $\mathrm{Ker}(N)$ is the set of elements $\sum_{s\in S} a_s s$ with $\sum_{g\in G} a_{g(s)} = 0$ (for some fixed $s \in S$), which is the set of elements with $\sum_{s\in S} a_s = 0$ because A has no $|G|$ torsion. But this space is obviously spanned by the set of elements of the form $s - g(s)$ as G acts transitively on S. This proves lemma 4.1.

Lemma 4.2. *The set of permutation modules is closed under taking sums, tensor products, and symmetric powers.*

Proof. The operations of taking sums or tensor products correspond to taking unions and products on the permutation representations. The operation of taking a symmetric n'th power of a representation corresponds to taking the set of "multisubsets" of cardinality n of a permutation representation S. (A multisubset is like a subset except that we are allowed to take several copies of some element.) This proves lemma 4.2.

Remark. If A is a permutation module for G coming from a permutation in which every element of G acts as an even permutation then all exterior powers of A are permutation modules for G. In particular if G has odd prime order then the set of permutation modules is closed under taking exterior powers, but this is false when G has order 2.

Lemma 4.3. *If A is a free $\mathbf{Z}_p$ module acted on by the group G generated by an element g of order p, then $\hat{H}^1(g, A) = 0$ if and only if A is a permutation module of G.*

Proof. If A is a permutation module then $\hat{H}^1(g, A)$ vanishes by lemma 4.1 (using the fact that $\hat{H}^{-1}$ is isomorphic to $\hat{H}^1$ for cyclic groups). Conversely if $\hat{H}^1(g, A)$ vanishes then A is a sum of the indecomposable representations $\mathbf{Z}_p$ and $\mathbf{Z}_p[G]$ (because the other indecomposable representation $I \otimes \mathbf{Z}_p$ has nonvanishing $\hat{H}^1$ by the remarks after lemma 2.1), and these are both permutation modules of G. This proves lemma 4.3.

Lemma 4.4. *The free modules over $\mathbf{Z}_p$ acted on by $G = \mathbf{Z}/p\mathbf{Z}$ with vanishing $\hat{H}^1$ are closed under taking sums, products, and symmetric powers.*

Proof. This follows immediately from lemmas 4.2 and 4.3.

Lemma 4.5. *If g is an element of odd prime order p in $M_{24} \subset \mathrm{Aut}(\Lambda)$ (where Λ is the Leech lattice) then $\hat{H}^1(g, \Lambda) = \hat{H}^1(g, \Lambda \otimes \mathbf{Z}_p) = 0$.*

Proof. This follows from lemmas 2.1 and 4.1 and the fact that as a representation of M_{24}, $\Lambda \otimes R$ is isomorphic to the representation coming from the usual permutation representation of M_{24} whenever R contains $1/2$, and in particular whenever R is the ring of p-adic integers for p an odd prime. (Recall that the Leech lattice Λ contains a set of 24 pairwise orthogonal vectors of squared length 8 which are acted on by M_{24} as the standard permutation of M_{24} and which generate the $Z[1/2]$-module $\Lambda \otimes Z[1/2]$.) This proves lemma 4.5.

Theorem 4.6. *Suppose g is an element of odd prime order p in $\mathrm{Aut}(\Lambda)$ such that $\hat{H}^1(g, \Lambda) = 0$. Then $\hat{H}^1(g, V_\Lambda) = 0$.*

Proof. Recall from section 3 that there is an integral form V_Λ of the vertex algebra of Λ, which is graded by Λ.

We will show that the sum of the degree-λ pieces $V_{\Lambda,\lambda}$ of V_λ for λ running over the elements of some orbit of G on Λ has vanishing first cohomology, which will prove theorem 4.6. If λ is not fixed by g then the sum of the spaces $V_{\Lambda,g^i(\lambda)}$ is a free module over the group ring of G and so all its cohomology groups vanish. Hence we may assume that λ is fixed by g. In this case, $V_{\Lambda,\lambda}$ is isomorphic to the degree 0 piece $V_{\Lambda,0}$ as a G module (under multiplication by e^λ). (At this point we are implicitly using the fact that g has odd order to deduce that g fixes e^λ whenever g fixes λ; if g has even order we need to be more careful.) Hence we have reduced the proof of theorem 4.6 to showing that $\hat{H}^1(g, V_{\Lambda,0}) = 0$, or equivalently to showing that $\hat{H}^1(g, V_{\Lambda,0} \otimes \mathbf{Z}_p) = 0$.

By assumption on g and by lemma 4.3 we can choose a basis $a_1, \ldots, a_{24}$ of $\Lambda \otimes \mathbf{Z}_p$ such that the action of g on $\Lambda \otimes \mathbf{Z}_p$ is given by some permutation representation on this basis. The ring $V_{\Lambda,0}$ is the polynomial ring over $\mathbf{Z}_p$ generated by the elements $e^{-a_j} D^{(i)}(e^{a_j})$ for $i \geq 1$, $1 \leq j \leq 24$. This polynomial ring, considered as a representation of G over $\mathbf{Z}_p$, is a direct sum of products of symmetric powers of $\Lambda \otimes \mathbf{Z}_p$, so by lemma 4.4 its first cohomology vanishes. This proves theorem 4.6. (Warning: $V_{\Lambda,0}$ is strictly larger than the polynomial ring generated by the elements $D^{(i)}(a_j)$ for $i \geq 1$, $1 \leq j \leq 24$.)

Theorem 4.7. *Suppose g is an element of the monster of type 3A, 3C, 5A, 7A, 11A, or 23A, and let $V[1/2]$ be the $\mathbf{Z}[1/2]$-form of the monster vertex algebra constructed in 3.2. Then $\hat{H}^1(g, V[1/2]) = 0$.*

Proof. The element g in the monster commutes with an element h of type 2B because it can be seen from the monster character table in [C] that the monster always has an element of order $2p$ (and type 6C, 6F, 10B, 14B, 22B, or 46A) whose square is in the conjugacy class pA or 3C of g and whose p'th power is of type 2B. The centralizer of h is a group with structure $2^{1+24}.Co_1$, which contains a split extension of M_{24} by a 2-group of the form $2^{1+24}.2^{12}$. We can identify g with an element of this M_{24} group.

In particular we can consider g to be an element of $\mathrm{Aut}(\hat{\Lambda})$, so that it acts on the integral form V_Λ of the vertex algebra of the Leech lattice Λ. The 1st cohomology of g with coefficients in any direct summand of $V_\Lambda \otimes \mathbf{Z}[1/2]$ must also vanish, as it is a direct summand of $\hat{H}^1(g, V_\Lambda[1/2]) = 0$. By theorem 3.2, $V[1/2]$ is a direct sum of modules (V^{00}, V^{01}, V^{10}, and V^{11}) which are isomorphic as modules over g to direct summands of $V_\Lambda[1/2]$, so $\hat{H}^1(g, V[1/2]) = 0$. This proves theorem 4.7.

Corollary 4.8. *(The modular moonshine conjectures [R94] for the primes 3, 5, 7, 11, 23.) Suppose g is an element of order p in one of the conjugacy classes 3A, 3C, 5A, 7A, 11A, or 23A of the monster. Then there is a vertex algebra ${}^gV = \oplus_n {}^gV_n$ defined over the field with p elements with the following properties.*

1 *If g is not in the class 3C then ${}^gV_2$ contains a conformal vector (generating an action of the mod p Virasoro algebra) and has a nonsingular bilinear form compatible with the vertex algebra structure.*

2 *${}^gV$ is acted on by the group $C_M(g)/\langle g\rangle$, and this action preserves the conformal vector and the bilinear form.*

3 *The Brauer trace $\sum_n \mathrm{Tr}(h|{}^gV_n)q^n$ of any p-regular element h of $C_M(g)/\langle g\rangle$ on ${}^pV$ is a Hauptmodul, and is equal to the Hauptmodul of the element gh of the monster.*

Everything in corollary 4.8 except for the existence of a conformal vector follows from 4.7 and other things stated above. To show the existence of a conformal vector it is sufficient to show that the image of the conformal vector ω in $V[1/2]$ is nonzero, and to do this it is sufficient to find some vector fixed by g that has nonzero inner product with $\omega(\mathrm{mod} p)$. We let $a_1, \dots a_{24}$ be an orthonormal base of $\Lambda \otimes \mathbf{Z}_p$ corresponding to a set acted on by M_{24}, and the element g fixes at least one of the a_i's (as g is not of type 3C), say a_1. Then the conformal vector ω is $\sum a_i(1)a_i(1)/2$ which has nonzero inner product 1 with the element $a_1(1)a_1(1)$ $\mathrm{mod} p$. This proves 4.8.

The head representation $\hat{H}^0(g, V[1/2]_1)$ containing the conformal vector ω splits slightly differently depending on whether $p > 3$ or $p \leq 3$. If $p > 3$ then ω has norm $12 \bmod p$, which is nonzero so that $\hat{H}^0(g, V[1/2]_1)$ splits as the orthogonal direct sum of a 1-dimensional space spanned by ω and its orthogonal complement (and this splitting is obviously invariant under $C_M(g)$). If $p = 3$ and g is an element of type $3A$ then ω has norm $12 \equiv 0 \bmod 3$ so it is contained in its orthogonal complement, and $\hat{H}^0(g, V[1/2]_1)$ has a composition series of the form 1.781.1 (as a $C_M(g)/\langle g\rangle = Fi'_{24}$ module). The Atlas [C] states that the 781 dimensional module has an algebra structure mod 3, but this seems to be a mistake and the construction there only gives an algebra structure on 1.781.1 mod 3.

When g is in the class 3C then the vertex algebra $\hat{H}^0(g, V[1/2])$ does not have a conformal vector, and in fact not only is the image of ω equal to zero, but the whole degree 1 space $\hat{H}^0(g, V[1/2]_1)$ is zero. However we can turn $\hat{H}^0(g, V[1/2])$ into a better vertex algebra by "compressing" it. If we look at the series $q^{-1} + 248q^2 + 4124q^5 + \cdots$ we see that only every 3rd term of the graded space $\hat{H}^0(g, V[1/2])$ is nonzero. Hence we change the grading of the piece of degree $3n-1$ to n. We also change the vertex operator $v(x)$ of $v \in \hat{H}^0(g, V[1/2])$ to $v(x^{1/3})$. It is easy to check that this defines a new vertex algebra structure on $\hat{H}^0(g, V[1/2])$, (because if we are in characteristic 3 then $x - y = (x^{1/3} - y^{1/3})^3$), whose homogeneous degrees are the coefficients of $1 + 248q + 4124q^2 + \cdots$. The compression of the vertex algebra $\hat{H}^0(g, V[1/2])$ is probably isomorphic to the vertex algebra $V_{E_8}/3V_{E_8}$ (the vertex algebra of the E_8 lattice reduced mod 3) which is acted on by the finite group $E_8(\mathbf{F_3})$, because both vertex algebras have the same graded dimension and are both acted on by the Thompson group. (It is well known that the Thompson group $C_M(g)/\langle g\rangle$ is contained in $E_8(\mathbf{F_3})$.) Warning: the vertex algebra $V_{E_8}/3V_{E_8}$ can be lifted to characteristic 0, but the the action of the Thompson group does not lift to an action on the vertex algebra V_{E_8}.

5. The case $p = 2$.

In this section we will extend the proof of the modular moonshine conjectures (Corollary 4.8) to cover the case of the elements of type 2A in the monster, to obtain a vertex algebra over $\mathbf{F_2}$ acted on by the baby monster. The proof is similar to the cases for elements of odd order, except that there are several extra technical complications, and we have to use one assertion which we have not completely proved.

The theorems in this section depend on the following assumption about the monster vertex algebra. We have not rigorously verified this, but we sketch how its verification should go.

Assumption 5.1. *There is a $\mathbf{Z}[1/3]$-form $V[1/3]$ of the monster vertex algebra with a self dual bilinear form such that $V[1/3]$ can be written as a direct sum of $2.M_{12}.2$ modules each of which is isomorphic to a submodule of the $\mathbf{Z}[1/3]$-form $V_\Lambda[1/3]$ of V_Λ.*

We outline a possible proof of 5.1. (Roughly speaking, this outline would become a correct proof if Dong, Mason or Montague do not need to divide by any integer other than 3 in their constructions.) We obviously cannot use the FLM construction of the monster vertex algebra, which writes V as the sum of the eigenspaces of an element of type 2B, because this involves inverting 2 and gives a 2-divisible module whose cohomology vanishes. Instead we use the construction of the monster vertex algebra as a sum of eigenspaces of an element of type 3B, due independently to Dong and Mason [DM] and Montague [M]. This should produce a $\mathbf{Z}[1/3,\omega]$-form $V[1/3,\omega]$ of the monster vertex algebra as follows, where ω is a cube root of 1 satisfying $\omega^2+\omega+1=0$. We first note that modules over $\mathbf{Z}[1/3,\omega]$ can be written as the sums of eigenspaces of any group $\langle h\rangle$ of order 3 acting on them, because $a=\frac{a+h(a)+h^2(a)}{3}+\frac{a+\omega h(a)+\omega^2h^2(a)}{3}+\frac{a+\omega^2h(a)+\omega h^2(a)}{3}$. The vertex algebra $V[1/3,\omega]$ is a sum of 3 submodules V^0, V^1, and V^2, which are the eigenspaces of an element h of type 3B. The space V^0 is isomorphic to the subspace of $V_\Lambda[1/3,\omega]$ fixed by an element of order 3 coming from a fixed point free element of order 3 of $\mathrm{Aut}(\Lambda)$, and this defines the $\mathbf{Z}[1/3,\omega]$-form of V^0. The spaces V^i each split into the sum of 3 subspaces V^{ij} which are the eigenspaces of a group $(\mathbf{Z}/3\mathbf{Z})^2$ in the monster, and Dong and Mason and Montague show that there is a group $SL_2(\mathbf{F_3})$ acting transitively on the spaces V^{ij} for $(i,j)\neq(0,0)$. We use this group to transport the $\mathbf{Z}[1/3,\omega]$-form from V^{01} and V^{02} to V^{ij} for $i\neq 0$. This defines the $\mathbf{Z}[1/3,\omega]$-form on the monster vertex algebra (and it also has a self dual $\mathbf{Z}[1/3,\omega]$ bilinear form and a conformal vector, coming from the same structures on V_Λ).

We now reduce this $\mathbf{Z}[1/3,\omega]$-form to a $\mathbf{Z}[1/3]$-form. We have an operation of complex conjugation on V^0 coming from complex conjugation on $V_\Lambda\otimes\mathbf{Z}[1/3,\omega]$, and the action of $SL_2(\mathbf{F_3})$ can be used to transfer this to complex conjugation on all the spaces $V^{ij}\oplus V^{2i,2j}$, and hence to $V[1/3,\omega]$. This conjugation preserves the vertex algebra structure and commutes with the monster, and satisfies $\overline{\omega v}=\bar\omega\bar v$. Any element of $V[1/3,\omega]$ can be written uniquely in the form $a+\omega b$ with a and b fixed by conjugation, because finding a and b involves solving the 2 linear equations $v=a+b\omega$, $\bar v=a+\bar\omega v$, and there is a unique solution for a and b because the determinant of this system of equations is $\omega-\bar\omega$ which is a unit in $\mathbf{Z}[1/3,\omega]$. Hence if we define $V[1/3]$ to be the fixed points of the conjugation, we see that $V[1/3]\otimes\mathbf{Z}[1/3,\omega]=V[1/3,\omega]$. Hence $V[1/3]$ is a $\mathbf{Z}[1/3]$-form for the monster vertex algebra. This completes the arguments for assumption 5.1.

Theorem 5.2. *If we assume that 5.1 is correct then there is a vertex algebra defined over $\mathbf{F_2}$ acted on by the baby monster, such that the McKay-Thompson series of every element of the baby monster is a Hauptmodul. The vertex algebra has a conformal vector and a compatible self dual bilinear form, also invariant under the baby monster.*

Proof. We will only give details for parts of the argument that differ significantly from the proof of corollary 4.8. We choose an element $g\in M_{24}\subset\mathrm{Aut}(\Lambda)$ of order 2 and trace 8. (There is a unique conjugacy class of such elements in M_{24}.)

The first step is to prove that $\hat H^1(g,\Lambda)=\Lambda_{\langle g\rangle}/\mathrm{Ker}(N)=0$, which we do by direct calculation as follows. We can assume that g acts on Λ in the usual coordinates by acting as -1 on the first 8 coordinates and as $+1$ on the last 16. Then $\mathrm{Ker}(N)$ is the sublattice of Λ of vectors whose last 16 coordinates vanish. This lattice is a copy of the E_8 lattice with norms doubled (so it has 240 vectors of norm 4 and so on). It is generated by its norm 4 vectors and the centralizer of g acts transitively on these norm 4 vectors, so to prove that $\Lambda_{\langle g\rangle}=\mathrm{Ker}(N)$ we only need to prove that one of these norm 4 vectors is of the form $v-g(v)$ for some vector $v\in\Lambda$. We can do this by taking v to be the vector $(1^{23},-3)$. Therefore $\hat H^1(g,\Lambda)=0$.

From this it follows as in theorem 4.6 that $\hat H^1(\hat g,V_\Lambda)=0$, provided we show that g lifts to an element in $\mathrm{Aut}(\hat\Lambda)$ which we denote by $\hat g$, such that $\hat g$ also has order 2 and $\hat g$ fixes every element e^a of the twisted group ring of Λ corresponding to an element a of Λ fixed by g. (If g is of odd order the existence of a good element $\hat g$ is trivial, but for some elements of even order, for example those of trace 0, a lift with these properties does not exist.) Lemma 12.1 of [B92] says that such a lift $\hat g$ exists provided that $(v,g(v))$ is even for all elements $v\in\Lambda$, which follows because $(v,g(v))=((v-g(v),v-g(v))-(v,v)-(g(v),g(v)))/2$, and $(v,v)=(g(v),g(v))$ is even while $v-g(v)\in\mathrm{Ker}(N)$ has norm divisible by 4. Hence $\hat H^1(\hat g,V_\Lambda)=0$.

In particular any submodule of $V_\Lambda[1/3]$ invariant under $\hat g$ also has vanishing $\hat H^1$, so by assumption 5.1 $\hat H^1(\hat g,V[1/3])=0$. We can now follow the argument proving corollary 4.8 to construct a vertex algebra satisfying the conditions of theorem 5.2. This proves theorem 5.2.

Remark. In the proof of theorem 5.2 we should really have checked that the 2 vertex algebras constructed by Frenkel-Lepowsky-Meurman and by Dong-Mason-Montague are isomorphic as modules over the monster, so that the traces of elements of the monster are Hauptmoduls. Dong and Mason have announced [DM] that these two vertex algebras are indeed isomorphic as vertex algebras acted on by the monster. It would also be easy to show that they are isomorphic as modules over the monster (but not as vertex algebras) by using the results of [B92] to calculate the character of the Dong-Mason-Montague algebra.

It seems likely that $V[1/2]$ and $V[1/3]$ both come from a conjectural integral form V of the monster vertex algebra by tensoring with $\mathbf{Z}[1/2]$ or $\mathbf{Z}[1/3]$; see lemma 7.1.

If we look at the McKay-Thompson series for the element 2B of the monster we see that the coefficients are alternating, which suggests that $\hat{H}^0$ vanishes for half the homogeneous pieces of V, and $\hat{H}^1$ vanishes for the other half. The next theorem shows that this is indeed the case.

Theorem 5.3. *Assume that 5.1 is correct. If $\hat{g}$ is an element of type 2B in the monster, then $\hat{H}^0(\hat{g}, V[1/3]_n)$ vanishes if n is even, and $\hat{H}^1(\hat{g}, V[1/3]_n)$ vanishes if n is odd.*

Proof. By using the argument in the last few paragraphs of the proof of theorem 5.2 we see that it is sufficient to prove the same result for the cohomology with coefficients in V_Λ for a suitable element $\hat{g}$ of $\mathrm{Aut}(\hat{\Lambda})$. We will take g to be the automorphism -1 of Λ, and take its lift $\hat{g}$ in $\mathrm{Aut}(\hat{\Lambda})$ to be the element taking e^a to $(-1)^{(a,a)/2}(e^a)^{-1}$ in the group ring. It is easy to check that $\hat{g}$ is an automorphism of order 2. As a graded $\hat{g}$ module, V_Λ is the tensor product of the twisted group ring and the module $V_{\Lambda,0}$ as in theorem 4.6. The twisted group ring is a sum of infinitely many free modules over the group ring and one copy of $\mathbf{Z}$ (spanned by 1), so its 0'th cohomology group is $\mathbf{Z}/2\mathbf{Z}$, and its first cohomology vanishes. Hence $H^*(\hat{g}, V_\Lambda) = H^*(\hat{g}, V_{\Lambda,0})$. We let ω be the involution of $V_{\Lambda,0}$ that is the product of $\hat{g}$ and the involution acting as -1^n on the piece of degree n. As multiplying the action of $\hat{g}$ by -1 exchanges H^0 and H^1, we see that to prove theorem 5.3 we have to show that $H^1(\omega, V_{\Lambda,0}) = 0$. To do this it is sufficient to show that $V_{\Lambda,0}$ is a permutation module for ω.

We now complete the proof of theorem 5.3 by finding an explicit basis of $V_{\Lambda,0}$ described in terms of S-functions, on which ω acts as a permutation. If we choose a basis $a_1, \dots, a_{24}$ of Λ, then $V_{\Lambda,0}$ is the polynomial ring generated by the elements $h_{i,n} = e^{-a_i}D^{(n)}e^{a_i}$, $1 \le i \le 24$, $n \ge 1$. The element $\hat{g}$ takes $h_{i,n}$ to $e^{a_i}D^{(n)}e^{-a_i}$, so $\omega(h_{i,n}) = e_{i,n}$, where $e_{i,n} = (-1)^n e^{a_i}D^{(n)}e^{-a_i}$. From the relation

$$\sum_m D^{(m)}(e^{a_i})D^{(n-m)}(e^{-a_i}) = D^{(n)}(e^{a_i}e^{-a_i}) = 0$$

for $n \ge 1$ we find that

$$\sum_{0 \le m \le n} (-1)^m e_{i,m} h_{i,n-m} = 0$$

for $n \ge 1$, which recursively defines the $e_{i,n}$'s in terms of the $h_{i,n}$'s. If we identify the $h_{i,n}$'s with the n'th complete symmetric function (see [Mac], page 14) then the ring generated by the $h_{i,n}$'s for some fixed i is the ring of all symmetric functions. By the formulas 2.6' and 2.7 on page 14 of [Mac] we see that the $e_{i,n}$'s are then identified with the elementary symmetric functions, and ω with the canonical involution ω of [Mac]. The ring of all symmetric functions is spanned by the S-functions s_λ for permutations λ [Mac, p. 24, formulas 3.1 and 3.3], and the action of ω on the S-functions is given by

$$\omega(s_\lambda) = s_{\lambda'}$$

[Mac, p. 28, formula 3.8] where λ' is the conjugate partition of λ [Mac, p. 2]. Therefore the ring generated by the $h_{i,n}$'s for any fixed i is a permutation module for ω, and therefore so is $V_{\Lambda,0}$ because it is the tensor product of 24 permutation modules. This proves theorem 5.3.

Of course, as we are working in characteristic 2, there is really no difference between a vertex algebra and a vertex superalgebra, so the superalgebra associated to 2B can be considered as a vertex algebra with dimensions given by the coefficients of the series $q^{-1} + 276q + 2048q^2 + \cdots$. However it seems more natural to think of it as a superalgebra.

This superalgebra is acted on by the group $2^{24}.Co_1$. The next paragraph suggests that the 2^{24} may act trivially, so that we get an action of Co_1, but we have not proved this.

We can also ask if this superalgebra can be lifted to characteristic 0. It turns out that there is a vertex superalgebra (constructed below) with the right grading acted on by $\mathrm{Aut}(\Lambda)$. In fact this superalgebra has

2 different natural actions of Aut(Λ), one faithful, and one for which the center of Aut(Λ) acts trivially. Unfortunately we do not know of an integral form invariant under either action of Aut(Λ), but it seems reasonable to guess that both actions have invariant integral forms, and the reduction mod 2 of either integral form is the mod 2 vertex superalgebra above. (This requires that the 2^{24} acts trivially on this mod 2 vertex algebra.)

We can construct this superalgebra in characteristic 0 easily as the vertex superalgebra of the odd 12-dimensional unimodular lattice with no roots D_{12}, which is the set of vectors $(x_1, \ldots, x_{24})$ in $\mathbf{R}^{24}$ such that all the x_i's are integers or all are integers $+1/2$, and their sum is even. This vertex superalgebra is acted on by the spinor group $Spin_{12}(\mathbf{R})$. The group Aut(Λ) in $SO_{24}(\mathbf{R})$ can be lifted to the spin group as it has vanishing Schur multiplier, so we get an action of Aut(Λ) on the vertex algebra. The reason why we get two actions is that there are two conjugacy classes of embeddings of Aut(Λ) in $SO_{24}(\mathbf{R})$ (which are interchanged by a reflection). The spin group does not act faithfully because an element of order 2 in the center acts trivially, and it is not hard to check that for exactly one of the classes of embeddings of Aut(Λ) in $SO_{24}(\mathbf{R})$ the element $-1 \in \mathrm{Aut}(\Lambda)$ lifts to the element of $Spin_{24}(\mathbf{R})$ acting trivially.

The existence of these two actions of Aut(Λ) has the curious consequence that every second coefficient of $q^{-1} + 276q + 2048q^2 + \cdots$ is in a natural way the dimension of 2 different representations of Aut(Λ), in which the nontrivial element of the center acts as either $+1$ or -1. For example 2048 decomposes as either $1771 + 276 + 1$ or $2024 + 24$.

6. Example: the Held group.

We give some numerical tables to illustrate the case when g is an element of type 7A in the monster, with centralizer $\langle g\rangle \times He$. First we give the 7-modular character table of He (taken from [R88]), followed by the modular characters of the first few head characters of $\hat{H}^0(g, V[1/2])$ (which can be read off from [CN] table 4). The bottom line gives the corresponding conjugacy classes in the monster group (the one whose Hauptmodul is given by the coefficients of the head characters).

	$1A$	$2A$	$2B$	$3A$	$3B$	$4A$	$4B$	$4C$	$5A$	$6A$	$6B$	$8A$	$10A$	$12A$	$12B$	$15A$	$17A$	$17B$
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	50	10	2	5	−1	2	2	−2	0	1	−1	0	0	−1	−1	0	−1	−1
	153	9	−7	0	3	−3	1	1	3	0	−1	1	−1	0	1	0	0	0
	426	26	10	3	−3	−2	2	2	1	−1	1	−2	1	1	−1	−2	1	1
	798	38	14	6	3	6	−2	2	−2	2	−1	0	−2	0	1	1	−1	−1
	1029	−35	21	21	0	−7	−3	1	4	1	0	−1	0	−1	0	1	$\frac{1-\sqrt{17}}{2}$	$\frac{1+\sqrt{17}}{2}$
	1029	−35	21	21	0	−7	−3	1	4	1	0	−1	0	−1	0	1	$\frac{1+\sqrt{17}}{2}$	$\frac{1-\sqrt{17}}{2}$
	1072	16	−16	10	−2	0	0	0	−3	−2	2	0	1	0	0	0	1	1
	1700	20	4	−10	5	0	−4	−4	0	2	1	0	0	0	−1	0	0	0
	3654	−10	38	−9	3	2	−2	−2	4	−1	−1	2	0	−1	1	1	−1	−1
	4249	9	−7	−8	−5	−3	1	1	−1	0	−1	1	−1	0	1	2	−1	−1
	6154	−70	−22	−2	7	6	2	2	4	2	−1	−2	0	0	−1	−2	0	0
	6272	−64	0	35	8	−8	0	0	−3	−1	0	0	1	1	0	0	−1	−1
	7497	81	−7	0	3	−3	1	−3	−3	0	−1	−1	1	0	1	0	0	0
	13720	−56	56	−14	7	0	8	0	−5	−2	−1	0	−1	0	−1	1	1	1
	14553	9	−7	0	−9	9	−7	1	3	0	−1	1	−1	0	−1	0	1	1
	17493	21	21	−21	0	−7	−3	5	−7	3	0	1	1	−1	0	−1	0	0
	23324	−196	28	14	−7	0	4	−4	−1	2	1	0	−1	0	1	−1	0	0
H_{-1}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
H_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_1	51	11	3	6	0	3	3	−1	1	2	0	1	1	0	0	1	0	0
H_2	204	20	−4	6	3	0	4	0	4	2	−1	2	0	0	1	1	0	0
H_3	681	57	9	15	0	1	9	1	6	3	0	1	2	1	0	0	1	1
H_4	1956	92	−12	30	0	0	12	0	6	2	0	2	2	0	0	0	1	1
H_5	5135	207	15	41	8	7	15	−1	10	9	0	3	2	1	0	1	1	1
H_6	12360	312	−24	66	0	0	24	0	10	6	0	4	2	0	0	1	1	1
H_7	28119	623	39	111	0	7	39	3	19	11	0	5	3	1	0	1	1	1
H_8	60572	932	−52	146	11	0	52	0	22	14	−1	6	2	0	1	1	1	1
H_9	125682	1674	66	222	0	18	66	−2	32	18	0	8	4	0	0	2	1	1
H_{10}	251040	2464	−96	336	0	0	96	0	40	16	0	8	4	0	0	1	1	1
	$7A$	$14A$	$14B$	$21A$	$21C$	$28A$	$28B$	$28C$	$35A$	$42A$	$42C$	$56A$	$70A$	$84A$	$84C$	$105A$	$119A$	$119A$

The values for the first 50 head characters can be extracted from the tables in [MS], by looking up the values of the head characters in the monster conjugacy classes listed in the last line. The paper [MS] also gives the decompositions of the first 50 head characters of the monster into irreducibles, which can be compared with the next table. (The top left corners of both tables are very similar.)

The next table gives the decomposition of the first few head characters H_i of $\hat{H}^0(g, V[1/2])$ into irreducible characters. The columns correspond to the irreducible characters arranged in order of their degrees (which are given in the first row). For example the 5th row means that the composition factors of the head representation H_2 are the representations of dimension 1, 50, and 153.

	1	50	153	426	798	1029	1029	1072	1700	3654	4249	6154	6272	7497	13720	14553	17493	23324
H_{-1}	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_4	2	3	2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H_5	4	5	3	2	1	0	0	1	1	0	0	0	0	0	0	0	0	0
H_6	4	7	5	3	1	0	0	3	1	0	1	0	0	0	0	0	0	0
H_7	7	12	8	7	3	0	0	5	2	1	2	0	0	0	0	0	0	0
H_8	8	16	13	9	4	0	0	9	4	1	4	1	0	1	0	0	0	0
H_9	12	25	18	17	8	0	0	15	7	4	7	1	0	2	0	1	0	0
H_{10}	14	35	29	26	11	0	0	27	12	6	15	4	1	4	0	2	0	0
H_{11}	23	53	43	45	21	0	0	43	22	13	25	7	2	8	1	4	1	0
H_{12}	26	75	67	68	29	0	0	74	37	21	48	16	4	14	1	9	2	1
H_{13}	40	114	99	114	50	0	0	119	62	41	79	27	8	26	3	18	5	3
H_{14}	49	161	155	174	72	0	0	202	106	68	144	58	16	44	6	32	10	6
H_{15}	71	243	233	290	119	0	0	328	176	124	244	99	28	74	13	58	20	14
H_{16}	88	348	358	446	173	0	0	543	292	204	422	186	53	124	22	103	35	28
H_{17}	128	519	543	723	279	1	1	876	484	355	700	319	91	206	43	176	64	54
H_{18}	161	752	831	1121	408	1	1	1425	785	578	1179	562	160	335	72	300	110	100
H_{19}	231	1125	1263	1793	643	4	4	2280	1271	969	1927	937	270	542	129	500	190	180
H_{20}	298	1637	1932	2769	951	6	6	3656	2043	1560	3159	1591	457	868	215	824	315	313

Notice that the multiplicities of nontrivial representations of small dimension have a tendency to start off with some of the values of the series $1, 1, 2, 3, 5, 7, \ldots$ which are the values $p(n)$ of the partition function. The same is true for the multiplicities of representations of the monster in the monster vertex algebra, and in that case it can be explained using the Virasoro algebra, and the fact that the Verma modules for the Virasoro algebra with $c = 24$, $h > 0$ are irreducible and have graded dimensions $1, 1, 2, 3, 5, 7, \ldots, p(n), \ldots$. (The irreducible factor of the Verma module with $c = 24$, $h = 0$ has pieces of dimension $p(n) - p(n-1) = 1, 0, 1, 1, 2, 2, 4, 4, 7, \ldots$, which more or less accounts for the initial multiplicities of the trivial character.) For our modular vertex algebras this explanation does not work so well, because the Verma modules are usually reducible over finite fields. We can get a small amount of information by examining how the Verma modules decompose, but this does not seem to be enough to account for why the numbers in the table above are so similar to the numbers we get when we look at the monster vertex algebra. On the other hand, if g has large prime order then the corresponding numbers do not seem to be similar to those for the monster; for example, if g has order 71 then the dimensions of the H_i's start off $1, 0, 1, 1, 1, \ldots$, which cannot be the dimensions corresponding to any representation of the Virasoro algebra in characteristic 0.

Silly question: why do the representations of dimension 1029 appear so late in the head characters?

7. Open problems and conjectures.

1. Can the information about modular representations be used to calculate the $|g|$-modular character tables of $C_M(g)$? The mod 7 character table of He has already been worked out in [R88], so the next simplest case is the mod 5 character table of the Harada-Norton group HN. For example, by cutting up the mod 5 vertex algebra of HN using the mod 5 Virasoro algebra we find that HN has representations over $\mathbf{F_5}$ of dimensions 1, 133, 626 and 2451 (which are probably irreducible). Unfortunately it seems to be difficult to get useful information like this from the later head representations, because we run into the problem that Verma modules over the Virasoro algebra mod p are not irreducible.
2. Does the monster vertex algebra have an integral form V such that each homogeneous piece is self dual under the natural bilinear form? It is easy to construct some monster invariant integral form by taking some integral form and taking the intersections of its conjugates under the action of the monster, but this will be far too small. The following lemma shows that we are quite close to constructing such an integral form.

Lemma 7.1. *Suppose that the spaces* $V[1/2] \otimes \mathbf{Z}[1/6]$ *and* $V[1/3] \otimes \mathbf{Z}[1/6]$ *are isomorphic as vertex algebras* *acted on by the monster. Then there exists an integral form of the monster vertex algebra with a compatible self dual integral bilinear form.*

Dong and Mason [DM] have announced that these two vertex algebras are isomorphic over the complex numbers, but their description of the proof (which has not appeared yet) sounds as if it might be hard to carry out over $\mathbf{Z}[1/6]$.

Proof. We denote $V[1/2]\otimes\mathbf{Z}[1/6]$ by $V[1/6]$, so we can assume that $V[1/6]$ contains $V[1/2]$ and $V[1/3]$ as subalgebras, and we define V to be $V[1/2]\cap V[1/3]$. It is obvious that V is a $\mathbf{Z}$-form of the monster vertex algebra, and we just have to check that the bilinear form on V is self dual. The embeddings of $V[1/2]$ and $V[1/3]$ into $V[1/6]$ preserve the conformal vector (as this is the only degree 2 vector ω fixed by the monster such that the operator ω_1 multiplies every vector by its degree), so the embeddings preserve the action of the Virasoro algebra. The bilinear forms are determined by the grading and vertex algebra structure and the action of the Virasoro algebra, so the embeddings also preserve the bilinear forms on all 4 algebras. If we look at the embedding of V into $V[1/2]$ we see that the bilinear form on V is self dual over $\mathbf{Z}_p$ for any odd prime p (as the bilinear form on $V[1/2]$ is self dual over $\mathbf{Z}[1/2]$), and similarly if we look at the embedding into $V[1/3]$ we see that the bilinear form on V is self dual over $\mathbf{Z}_p$ for any $p\neq 3$. Hence the symmetric bilinear form on V is self dual over all rings of p-adic integers, and is therefore self dual over $\mathbf{Z}$. This proves lemma 7.1.

It may be possible to prove that the vertex algebras $V[1/2]$ and $V[1/3]$ are isomorphic over $\mathbf{Z}[1/6]$ either by carrying out the proof suggested in [DM] over $\mathbf{Z}[1/6]$, or by constructing the monster vertex algebra as a sum of eigenspaces of an element of the monster of type 6B. This element corresponds to a fixed point free element of order 6 in $\mathrm{Aut}(\Lambda)$ (of trace 12) whose cube and square are the elements of orders 2 and 3 in $\mathrm{Aut}(\Lambda)$ used to construct $V[1/2]$ and $V[1/3]$.

The integral form V would give integral forms for all the homogeneous spaces V_n, and in particular would give an integral form on the Griess algebra V_1. This cannot be the same as the integral form constructed by Conway and Norton in [C85], because the one in [C85] contains the element $\omega/2$ (which is denoted by 1 there). It seems possible that Conway and Norton's integral form is spanned by $\omega/2$ together with the elements of V_1 which have integral inner product with $\omega/2$. Notice that the bilinear form used in [C85] is half the bilinear form on V_1.

3. Assume the integral form V exists. Is $\hat{H}^1(g,V)$ zero whenever g is an element of the monster whose Hauptmodul has no negative coefficients? (Theorem 4.7 shows this for some elements of odd order.) If the coefficients of the Hauptmodul for g alternate in sign, do the groups $\hat{H}^0(g,V_{2n})$ and $\hat{H}^1(g,V_{2n+1})$ vanish? (Theorem 5.3 proves this for elements of type 2B.)
4. What happens if g is an element of composite order? For example, if we look at the Hauptmodul for an element of type 4B we see that it starts off $q^{-1}+52q+834q^3+4760q^5$, and the coefficient 52 of q^1 is the dimension of the Lie algebra F_4, and the centralizer of an element of type 4B is of the form $(4\times F_4(\mathbf{F_2})).2$. This suggests there should be a modular vertex algebra corresponding to elements of type 4B, whose "compression" should be the reduction mod 2 or 4 of a vertex algebra for F_4 defined over $\mathbf{Z}$, in the same way that the compression of the vertex algebra for elements of type 3C is the reduction mod 3 of an algebra for E_8. (See the end of section 4.)
5. We can construct Lie algebras and superalgebras which have much the same relation to our modular vertex algebras as the monster Lie algebra [B92] has to the monster vertex algebra. We do this by using the $\mathbf{Z}_p$-forms of the monster vertex algebra to put $\mathbf{Z}_p$-forms on the monster Lie algebra (with a self dual symmetric invariant bilinear form), and then take the Tate cohomology of this $\mathbf{Z}_p$-form of the monster Lie algebra, which by the comments at the beginning of section 2 produces a Lie superalgebra. (Of course if we have a good integral form of the monster Lie algebra we can use this directly and not worry about $\mathbf{Z}_p$-forms.) The Lie algebras and superalgebras we get are similar to generalized Kac-Moody algebras, except that they are over fields of characteristic p rather than characteristic 0: they have a root system, a Cartan subalgebra, an invariant nonsingular symmetric bilinear form, a Cartan involution, and a $\mathbf{Z}$-grading with finite dimensional homogeneous pieces. Their structure as $C_M(g)$ modules can be described as follows: they have a $\mathbf{Z}^2$-grading, such that the piece of degree (m,n) is isomorphic to ${}^gV_{mn}$ if $(m,n)\neq(0,0)$, and the piece of degree $(0,0)$ (the Cartan subalgebra) is 2-dimensional and acted on trivially by $C_M(g)$. This suggests that there should be some sort of theory of "generalized Kac-Moody algebras mod$p$", which could be applied to study these algebras. For example, we could ask for the

Lie algebra homology groups of the positive degree subalgebras (which in characteristic 0 is equivalent to asking for the simple roots). Notice that many of these modular Lie algebras cannot be obtained by reducing some integral form of a generalized Kac-Moody algebra mod$p$, because the denominator formula shows that some of the simple roots would then have negative multiplicity.

6. The modules $\hat{H}^*(g, V[1/2])$ are acted on not just by the centralizer of g but by the normalizer $N_M(g)$ of g. What are the traces of elements of the normalizer of g that are not in the centralizer? It may be possible to do this by extending proposition 2.2 to the case when the subgroup $\langle g\rangle$ is only normal and not central. (Notice that if $\hat{H}^1(g, V[1/2])$ is nonzero then elements of the normalizer do not necessarily preserve the algebra structure on $\hat{H}^*(g, V[1/2])$.)
7. Give a complete proof of assertion 5.1.
8. Can the modular vertex algebras ${}^gV$ be lifted to characteristic 0 in some way? The answer to the strong form of this question is usually "no": it is easy to check that it is usually impossible to lift ${}^gV$ to a vertex algebra in characteristic 0 that is acted on by $C_M(g)$. However Queen [Q] found strong evidence that ${}^gV$ could be lifted to some $C_M(g)$ representation in characteristic 0 (which cannot carry an invariant vertex algebra structure). The representations that Queen found evidence for are now called twisted sectors. Perhaps these representations have some sort of "twisted" vertex algebra structure, where the vertex operators $a(z)$ have branch points of order p at the origin. It might be possible to use some sort of analogue of Witt vectors for vertex algebras to construct these. (But there is one serious obstruction to any canonical way of lifting some ${}^gV$'s to characteristic 0: the automorphism group in characteristic 0 is sometimes a *nonsplit* central extension of the automorphism group in characteristic p.) On the other hand we have seen (in the remarks at the end of sections 4 and 5) that if g is of type 3C or 2B then ${}^gV$ can probably be lifted to a vertex algebra in characteristic 0, which is not acted on by $C_M(g)$. Dong, Li and Mason [DLM] have recently made some progress on this question by constructing a twisted sector that is probably a lift to characteristic zero of the space ${}^gV$ when g is of type $2A$.
9. If $p = 3, 5, 7$, or 13 and g is an element of type pB in the monster corresponding to the group $\Gamma_0(p)$ then the group $C_M(g)/O_p(C_M(g))$ contains an element of order 2 in its center. We conjecture that this element of order 2 acts as $+1$ on $\hat{H}^0(g, V[1/2])$ and as -1 on $\hat{H}^1(g, V[1/2])$. This would imply that the graded modular characters of both $\hat{H}^0$ and $\hat{H}^1$ can be expressed as a linear combination of 2 Hauptmoduls. These are the only elements of prime order in the monster not already covered by theorems 4.7, 5.2, 5.3, and question 3 above, so an affirmative answer to this question and question 3 would mean that we would have a complete description of the modular characters of both $\hat{H}^0(g, V)$ and $\hat{H}^1(g, V)$ for all elements g of prime order in the monster.

References.

[AW] M. F. Atiyah, C. T. C. Wall, Cohomology of finite groups, in "Algebraic number theory", editors J. W. S. Cassels and A. Fröhlich, Academic Press 1967.

[B86] R. E. Borcherds, Vertex algebras, Kac-Moody algebras, and the monster. Proc. Natl. Acad. Sci. USA. Vol. 83 (1986) 3068-3071.

[B92] R. E. Borcherds, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math. 109, 405-444 (1992).

[C] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, R. A. Wilson, Atlas of finite groups, Clarendon Press, Oxford, 1985.

[C85] J. H. Conway, A simple construction for the Fischer-Griess monster group, Invent. Math. 79 (1985) p. 513-540.

[CN] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[CR] C. W. Curtis and I. Reiner, "Methods of representation theory Vol 1", Wiley Interscience, 1981 and 1990.

[DLM] C. Dong, H. Li, G. Mason, Some twisted sectors for the moonshine module, to appear in Contemporary Math.

[DM] C. Dong and G. Mason, On the construction of the moonshine module as a $\mathbf{Z}_p$-orbifold, Santa Cruz preprint, 1992. To appear in: Proc. 1992 Joint Summer Research Conference on Conformal Field Theory, Topological Field theory and Quantum Groups, Mount Holyoke, 1992, Contemporary Math.

[FLM] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[Mac] I. G. Macdonald, “Symmetric functions and Hall polynomials”, Oxford University press, 1979.

[MS] J. McKay, H. Strauss, The q-series of monstrous moonshine and the decomposition of the head characters, Comm. in Alg. (1990) 18, 253-278.

[M] P. Montague, Third and Higher Order NFPA Twisted Constructions of Conformal Field Theories from Lattices, preprint, submitted to Nuc. Phys. B.

[N] S. P. Norton, Generalized Moonshine, Proc. Symp. Pure Math. 47 (1987) p. 208-209.

[Q] L. Queen, Some relations between finite groups, Lie groups, and modular functions, PhD thesis, Cambridge University, England, 1980.

[R88] A. J. E. Ryba, Calculation of the 7-modular characters of the Held group. J. Algebra, 117, 240-255, 1988.

[R94] A. J. E. Ryba, Modular Moonshine?, 1994 preprint.

Modular Moonshine III. *11 October 1994, corrected 18 Nov 1997*

Duke Math. J. 93 (1998), no. 1, 129–154.

Richard E. Borcherds,*
D.P.M.M.S., 16 Mill Lane, Cambridge, CB2 2SB, England.

Mathematics department, University of California at Berkeley, CA 94720-3840, U. S. A.

reb@dpmms.cam.ac.uk
www.dpmms.cam.ac.uk/˜reb

In this paper we complete the proof of Ryba's modular moonshine conjectures [R] that was started in [B-R]. We do this by applying Hodge theory to the cohomology of the monster Lie algebra over the ring of p-adic integers in order to calculate the Tate cohomology groups of elements of the monster acting on the monster vertex algebra.

Contents.

1. Introduction.
Notation.
2. Representations of $\mathbf{Z}_p[G]$.
3. The $\mathbf{Z}[1/2]$-form of the monster Lie algebra.
4. The modular moonshine conjectures for $p \geq 13$.
5. Calculation of some cohomology groups.
6. Elements of type $3B$, $5B$, $7B$, and $13B$.
7. Open problems and conjectures.

1. Introduction.

This paper is a continuation of the earlier paper [B-R] so we will only briefly recall results that are discussed there in more detail. In [B-R] we constructed a modular superalgebra for each element of prime order in the monster, and worked out the structure of this superalgebra for some elements in the monster. In this paper we work out the structure of this superalgebra for the remaining elements of prime order.

Ryba conjectured [R] that for each element of the monster of prime order p of type pA there is a vertex algebra ${}^gV$ defined over the finite field $\mathbf{F}_p$ and acted on by the centralizer $C_M(g)$ of g in the monster group M, with the property that the graded Brauer character $\mathrm{Tr}(h|{}^gV) = \sum_n \mathrm{Tr}(h|{}^gV_n)q^n$ is equal to the Hauptmodul $\mathrm{Tr}(gh|V) = \sum_n \mathrm{Tr}(g|V_n)q^n$ (where V is the graded vertex algebra acted on by the monster constructed by Frenkel, Lepowsky and Meurman [FLM]). In [B-R] the vertex superalgebra ${}^gV$ was defined for any element $g \in M$ of odd prime order to be the sum of the Tate cohomology groups $\hat{H}^0(g, V[1/2]) \oplus \hat{H}^1(g, V[1/2])$ for a suitable $\mathbf{Z}[1/2]$ form $V[1/2]$ of V, and it was shown that $\mathrm{Tr}(h|{}^gV) = \mathrm{Tr}(h|\hat{H}^0(g, V[1/2])) - \mathrm{Tr}(h|\hat{H}^1(g, V[1/2]))$ was equal to the Hauptmodul of $gh \in M$. Hence to prove the modular moonshine conjecture for an element g of type pA it is enough to prove that $\hat{H}^1(g, V[1/2]) = 0$. In [B-R] this was shown by explicit calculation for the elements of type pA for $p \leq 11$, using the fact that these elements commute with an element of type $2B$. For $p \geq 13$ this method does not work as these elements do not commute with an element of type $2B$. The first main theorem of this paper is theorem 4.1 which states that if g is an element of prime order $p \geq 13$ not of type $13B$ then $\hat{H}^1(g, V[1/2]) = 0$ (assuming a certain condition about the monster Lie algebra, whose proof should appear later). Hence Ryba's conjectures are proved for all elements of M of type pA. (Actually this is not quite correct, because the proof for $p = 2$ in [B-R] assumes an unproved technical hypothesis.) The proof we give fails for exactly the cases already proved in [B-R].

We can also ask what happens for the other elements of order less than 13. The cases of elements of types $2B$ or $3C$ are also treated in [B-R]: for type $3C$, $\hat{H}^1(g, V[1/2])$ is again zero, and for type $2B$, $\hat{H}^1(g, V[1/2])$ is zero in even degree and $\hat{H}^0(g, V[1/2])$ is zero in odd degree. This leaves the cases where g is of type $3B$, $5B$, $7B$, or $13B$, when the Tate cohomology groups $\hat{H}^0(g, V[1/2])$ and $\hat{H}^1(g, V[1/2])$ are usually both nonzero in each degree. In these cases the structure of the cohomology groups is determined

---

* Supported by NSF grant DMS-9401186 and by a Royal Society research professorship.

by our second main result, theorem 6.1, which states that the unique element σ of order 2 in the center of $C_M(g)/O_p(C_M(g))$ acts as 1 on $\hat{H}^0(g, V[1/2])$ and as -1 on $\hat{H}^1(g, V[1/2])$. This determines the modular characters of both cohomology groups because $\mathrm{Tr}(h|\hat{H}^0(g, V[1/2])) + \mathrm{Tr}(h|\hat{H}^1(g, V[1/2]))$ is then given by $\mathrm{Tr}(h\sigma|\hat{H}^0(g, V[1/2])) - \mathrm{Tr}(h\sigma|\hat{H}^1(g, V[1/2]))$ which is the Hauptmodul of the element $gh\sigma \in M$. Hence both $\mathrm{Tr}(h|\hat{H}^0(g, V[1/2]))$ and $\mathrm{Tr}(h|\hat{H}^1(g, V[1/2]))$ can be written explicitly as linear combinations of two Hauptmoduls for any p-regular element $h \in C_M(g)$. We get two new modular vertex superalgebras acted on by double covers of sporadic groups when g has type $3B$ or $5B$: a vertex superalgebra over $\mathbf{F}_3$ acted on by $2.Suz$, and a vertex superalgebra over $\mathbf{F}_5$ acted on by $2.HJ$. For elements of types $7B$ and $13B$ the vertex superalgebras we get are acted on by the double covers $2.A_7$ and $2.A_4$ of alternating groups.

In particular if the integral form V of the monster vertex algebra discussed in [B-R] exists and has the properties conjectured there then all the cohomology groups $\hat{H}^i(g, V)$ are now known for all elements g of prime order in M.

We now discuss the proofs of the two main theorems. We first quickly dispose of the second main theorem: if an element $g \in M$ of odd order commutes with an element of type $2B$ then it is easy to work out its action on the $\mathbf{Z}[1/2]$-form of the monster vertex algebra and hence the Tate cohomology groups $\hat{H}^*(g, V[1/2])$ can be calculated by brute force, which is what we do in sections 5 and 6 for the elements of types $3B$, $5B$, $7B$, and $13B$.

The elements g of prime order $p \geq 13$ do not commute with any elements of type $2B$ (except when g has type $13B$ or $23AB$) so we cannot use the method above. Instead we adapt the proof of the moonshine conjectures in [B]. By using a $\mathbf{Z}_p$-form of the monster Lie algebra rather than a $\mathbf{Q}$-form we can find some complicated relations between the coefficients of the series $\sum_n \mathrm{Tr}(h|\hat{H}^0(g, V[1/2]_n))q^n$ and $\sum_n \mathrm{Tr}(h|\hat{H}^1(g, V[1/2]_n))q^n$. The reason we get more information by using $\mathbf{Z}_p$-forms rather than $\mathbf{Q}_p$-forms is that the group ring $\mathbf{Z}_p[\mathbf{Z}/p\mathbf{Z}]$ has 3 indecomposable modules which are free over $\mathbf{Z}_p$, while the group ring $\mathbf{Q}_p[\mathbf{Z}/p\mathbf{Z}]$ only has 2 indecomposable modules which are free over $\mathbf{Q}_p$, and this extra indecomposable module provides the extra information. The relations between the coefficients we get are similar to the relations defining completely replicable functions, except that some of the relations defining completely replicable functions (about "$(2p-1)/p^2$" of them) are missing. If p is large (≥ 13) we show that these equations have only a finite number of solutions. We then use a computer to find all the solutions, and see that the only solutions imply that $\hat{H}^1(g, V[1/2]) = 0$ for all g of prime order at least 13 other than those of type $13B$. This proof could probably be made to work for some smaller values of p such as 7 and 11, but the difficulty of showing that there are no extra solutions for the equations of the coefficients increases rapidly as p decreases, because for smaller p there are more equations missing.

Y. Martin [Ma] has recently found a conceptual proof that any completely replicable function is a modular function. Cummins and Gannon [CG] have greatly generalized this result using different methods, and showed that the functions are Hauptmoduls. It seems possible that their methods could be extended to replace the computer calculations in section 4, although Cummins has told me that this would probably require adding in some "missing" relations corresponding to the case $(p, mn) \neq 0$ in proposition 3.4.

We summarize the results of this paper and of [B-R] about the vertex superalgebras ${}^gV$ for all elements g of prime order p in the monster. If g is of type pA, or pB for $p > 13$, or $3C$, then ${}^gV$ is just a vertex algebra and not a superalgebra. If $p \geq 13$ this is proved in this paper by an argument which works for any self dual form of the monster vertex algebra but which relies on computer calculations and on a so far unpublished argument about integral forms of the monster Lie algebra. For $p < 13$ (or $p = 23$) this is proved in [B-R] by calculating explicitly for a certain $\mathbf{Z}[1/2]$ form of the monster vertex algebra; this avoids computer calculations but only works for one particular form of the monster vertex algebra. Also, if $p = 2$ the calculation depends on an as yet unproved technical assumption about the Dong-Mason-Montague construction of the monster vertex algebra from an element of type $3B$ ([DM] or [M]). If g is of type pB for $2 \leq p \leq 13$ then the super part of ${}^gV$ does not vanish. If $p = 2$ then ${}^gV$ is calculated explicitly in [B-R] (again using the unproved technical assumption) and it turns out that its ordinary part vanishes in odd degrees and its super part vanishes in even degrees. If $3 \leq p \leq 13$ then ${}^gV$ is calculated in this paper, and we find that there is an element in $C_M(g)$ acting as -1 on the super part and as 1 on the ordinary part of ${}^gV$. In all cases we have explicitly described the modular characters of $C_M(g)$ acting on the ordinary or super part of any degree piece of ${}^gV$ in terms of the coefficients of certain Hauptmoduls.

Notation.

$[A]$	is the element of K represented by the module A.
A, B, C	G-modules.
A_n	The alternating group on n symbols.
Aut	The automorphism group of something.
$c_g^+(n)$	The n'th coefficient of the Hauptmodul of $g \in M$, equal to $\mathrm{Tr}(1\|^g V_n) = \dim(\hat{H}^0(g, V[1/2]_n)) - \dim(\hat{H}^1(g, V[1/2]_n))$.
$c_g^-(n)$	$\dim(\hat{H}^0(g, V[1/2]_n)) + \dim(\hat{H}^1(g, V[1/2]_n))$.
$c_{m,n}$	Defined in proposition 4.3.
$\mathbf{C}$	The complex numbers.
$C_M(g)$	The centralizer of g in the group M.
Co_1	Conway's largest sporadic simple group.
E	The positive subalgebra of the monster Lie algebra.
F	The negative subalgebra of the monster Lie algebra.
$\mathbf{F}_p$	The finite field with p elements.
f	A homomorphism from K to $\mathbf{Q}$ defined in lemma 2.3.
g	An element of G, usually of order p.
g_i	The element g^i of G, used when it is necessary to distinguish the multiplication in the group ring from some other multiplication.
$\langle g \rangle$	The group generated by g.
G	A group, often cyclic of prime order p and generated by g.
$h(A), h_n$	See section 5.
H	The Cartan subalgebra of the monster Lie algebra.
$\hat{H}^i(G, A)$	A Tate cohomology group of the finite group G with coefficients in the G-module A.
$\hat{H}^i(g, A)$	means $\hat{H}^i(\langle g \rangle, A)$, where $\langle g \rangle$ is the cyclic group generated by g.
$\hat{H}^*(g, A)$	The sum of the Tate cohomology groups $\hat{H}^0(g, A)$ and $\hat{H}^1(g, A)$, considered as a super module.
HJ	The Hall-Janko sporadic simple group (sometimes denoted J_2).
I	The indecomposable module of dimension $p-1$ over $\mathbf{Z}_p[G]$, isomorphic to the kernel of the natural map from $\mathbf{Z}_p[G]$ to $\mathbf{Z}_p$ and to the quotient $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$.
Im	The image of a map.
K	A ring which is a free $\mathbf{Q}$-module with a basis of 3 elements $[\mathbf{Z}_p] = 1$, $[\mathbf{Z}_p[G]]$, and $[I]$.
Ker	The kernel of a map.
$\Lambda, \hat{\Lambda}$	The Leech lattice and a double cover of the Leech lattice.
$\Lambda^n(A)$	The n'th exterior power of A.
$\Lambda^*(A)$	The exterior algebra $\oplus_n \Lambda^n(A)$ of A.
L	An even lattice.
$\mathbf{m}$	The monster Lie algebra.
M	The monster simple group.
N_G	The element $\sum_{g \in G} g$ of $\mathbf{Z}_p[G]$.
$O_p(G)$	The largest normal p-subgroup of the finite group G.
p	A prime, usually the order of g.
$\mathbf{Q}, \mathbf{Q}_p$	The rational numbers and the field of p-adic numbers.
ρ	The Weyl vector of the monster Lie algebra.
$\mathbf{R}$	The real numbers.
R_p	A finite extension of the p-adic integers.
$R(i)$	Defined just before lemma 3.2.
σ	An automorphism of type $2B$ in the monster or the automorphism -1 of the Leech lattice.
S_n	A symmetric group.
$S^n(A)$	The n'th symmetric power of A.
$S^*(A)$	The symmetric algebra $\oplus_n S^n(A)$ of A.
Suz	Suzuki's sporadic simple group.
Tr	$\mathrm{Tr}(g\|A)$ is the usual trace of g on a module A if A is a module over a ring of characteristic 0, and the Brauer trace if A is a module over a field of finite characteristic.
$V[1/n]$	A $\mathbf{Z}[1/n]$-form of the monster vertex algebra.

V_Λ *The integral form of the vertex algebra of* $\hat{\Lambda}$.
V_n *The degree* n *piece of* V.
V^n *An eigenspace of some group acting on* V.
${}^gV$ *A modular vertex algebra or superalgebra given by* $\hat{H}^*(g, V[1/n])$ *for some* n *coprime to* $|g|$.
$\mathbf{Z}$ *The integers.*
$\mathbf{Z}_p$ *The ring of* p*-adic integers.*
ω *A cube root of 1 or a conformal vector.*
Ω *The Laplace operator on* $\Lambda^*(E)$.

2. Representations of $\mathbf{Z}_p[G]$.

We give some auxiliary results about modules over $\mathbf{Z}_p[G]$ where G is a cyclic group generated by g of prime order p, and in particular calculate the exterior and symmetric algebras of all indecomposable modules. All modules will be free over $\mathbf{Z}_p$ and will be either finitely generated or $\mathbf{Z}$ graded with finitely generated pieces of each degree. All tensor products will be taken over $\mathbf{Z}_p$.

Recall from [B-R section 2] that there are 3 indecomposable modules over $\mathbf{Z}_p[G]$, which are $\mathbf{Z}_p$, the group ring $\mathbf{Z}_p[G]$, and the module I of $\mathbf{Z}_p[G]$ that is isomorphic to the kernel of the natural map from $\mathbf{Z}_p[G]$ to $\mathbf{Z}_p$ and to the quotient $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$ (where $N_G = 1+g+g^2+\cdots+g^{p-1}$). Their Tate cohomology groups are given by $\hat{H}^0(g,\mathbf{Z}_p) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g,\mathbf{Z}_p) = 0$, $\hat{H}^0(g,\mathbf{Z}_p[G]) = 0$, $\hat{H}^1(g,\mathbf{Z}_p[G]) = 0$, $\hat{H}^0(g,I) = 0$, and $\hat{H}^1(g,I) = \mathbf{Z}/p\mathbf{Z}$.

Lemma 2.1. *The tensor products of these modules are given as follows:*

$$\begin{aligned}\mathbf{Z}_p \otimes X &= X \ \textit{(for any } X\textit{)}\\ \mathbf{Z}_p[G] \otimes X &= \mathbf{Z}_p[G]^{\dim(X)} \ \textit{(the sum of } \dim(X) \textit{ copies of } \mathbf{Z}_p[G]\textit{)}\\ I \otimes I &= \mathbf{Z}_p[G]^{p-2} \oplus \mathbf{Z}_p.\end{aligned}$$

Proof. The case $\mathbf{Z}_p \otimes X$ is trivial. If X has a basis $x_1, \ldots, x_n$ then for any fixed i the elements $g^j \otimes g^j(x_i)$ ($0 \le j < p$) form a basis for a submodule of $\mathbf{Z}_p[G] \otimes X$ isomorphic to $\mathbf{Z}_p[G]$, and $\mathbf{Z}_p[G] \otimes X$ is the direct sum of these submodules for all i, which proves the result about $\mathbf{Z}_p[G] \otimes X$. For the case $I \otimes I$ we look at the cohomology sequence of the short exact sequence

$$0 \to I \otimes I \to I \otimes \mathbf{Z}_p[G] \to I \otimes \mathbf{Z}_p \to 0$$

to see that $\hat{H}^0(I \otimes I) = \mathbf{Z}/p\mathbf{Z}$ and $\hat{H}^1(I \otimes I) = 0$, which implies that $I \otimes I$ must be the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$. The number of copies can be worked out by looking at the dimensions of both sides. This proves lemma 2.1.

Corollary 2.2. *If* A *and* B *are* $\mathbf{Z}_p[G]$*-modules then* $\hat{H}^*(g, A \otimes B) = \hat{H}^*(g, A) \otimes \hat{H}^*(g, B)$.

Proof. Check each of the 9 cases when A and B are indecomposable modules using lemma 2.1. This proves corollary 2.2.

We form the ring K which is a free $\mathbf{Q}$-module with a basis of 3 elements $[\mathbf{Z}_p] = 1$, $[\mathbf{Z}_p[G]]$, and $[I]$ corresponding to the indecomposable modules of $\mathbf{Z}_p[G]$, whose addition and multiplication are those induced by tensor products and direct sums of modules. If A is a $\mathbf{Z}_p[G]$-module then we write $[A]$ for the element of K represented by A. Notice that this is not the same as the (tensor product with $\mathbf{Q}$ of the) Grothendieck ring of finitely generated projective $\mathbf{Z}_p[G]$ modules because we do not assume the relations given by nonsplit short exact sequences. The Grothendieck ring (tensored with $\mathbf{Q}$) is a quotient of our ring K by the ideal generated by $[\mathbf{Z}_p] + [I] - [\mathbf{Z}_p[G]]$.

Lemma 2.3. *There are exactly 3 ring homomorphisms* $\dim$, $\mathrm{Tr}(g|.)$, *and* f *from* K *to* $\mathbf{Q}$ *(which correspond to the 3 elements of the spectrum of* K*). They are given by*

1. $\dim([\mathbf{Z}_p]) = 1$, $\dim([\mathbf{Z}_p[G]]) = p$, $\dim([I]) = p-1$.
2. $\mathrm{Tr}(g|[\mathbf{Z}_p]) = 1$, $\mathrm{Tr}(g|[\mathbf{Z}_p[G]]) = 0$, $\mathrm{Tr}(g|[I]) = -1$.
3. $f([\mathbf{Z}_p]) = 1$, $f([\mathbf{Z}_p[G]]) = 0$, $f([I]) = 1$.

Proof. From corollary 2.2 we know that the products in the ring K are given by $[\mathbf{Z}_p[G]][\mathbf{Z}_p[G]] = p[\mathbf{Z}_p[G]]$, $[\mathbf{Z}_p[G]][I] = (p-1)[\mathbf{Z}_p[G]]$, and $[I][I] = (p-2)[\mathbf{Z}_p[G]] + 1$. Hence if $x \in \mathbf{Q}$ and $y \in \mathbf{Q}$ are the images of $[\mathbf{Z}_p[G]]$ and I under some homomorphism they must satisfy $x^2 = px$, $xy = (p-1)x$, and $y^2 = (p-2)x + 1$. The only solutions to these equations are $(x, y) = (p, p-1)$, $(0, 1)$, or $(0, -1)$. This proves lemma 2.3.

The first two homomorphisms in lemma 2.3 are the homomorphisms which factor through the Grothendieck ring and will not be very interesting to us. We can use them in place of f in the arguments in the rest of the paper, but this only leads to relations between the coefficients of $\sum_n \dim(V_n)q^n$ and $\sum_n \mathrm{Tr}(g|V_n)q^n$, which are essentially the relations already used in [B] to determine these functions. It is the third homomorphism f which detects the integral structure on $V[1/2]$ which will give us enough extra information to prove theorem 4.1.

In general, taking the trace of any fixed element of a group gives a homomorphism from a representation ring of a group to some field, and quite often (e.g., for the complex representation ring of a group) this gives a 1:1 correspondence between homomorphisms and (certain sorts of) conjugacy classes of the group. Lemma 2.3 gives an example of a homomorphism f not constructed in this way.

A permutation module is a module for G with a basis acted on by G. We recall from [B-R] that a module A is a permutation module if and only if $\hat{H}^1(g, A) = 0$, which in turn is equivalent to saying that A is a sum of copies of $\mathbf{Z}_p$ and $\mathbf{Z}_p[G]$.

Lemma 2.4. *If A is a permutation module over any ring for any finite group g then so are the symmetric powers $S^n(A)$. If all elements of G act on a basis of A as products of cycles of odd length then the exterior powers $\Lambda^n(A)$ are also permutation modules.*

Proof. For symmetric powers this is proved in [B-R]. For the proof for exterior powers, let $a_1, \ldots, a_n$ be a basis for A acted on by G. Then G acts on the set of all elements of the form $\pm a_{i_1} \wedge \cdots \wedge a_{i_n}$, which consists of all elements of a basis together with their negatives. If the action of G on the basis of A has the property that all elements of G are products of cycles of odd length, then there is no element of G that maps any element of the form $\pm a_{i_1} \wedge \cdots \wedge a_{i_n}$ to its negative. This means that we can find a G-invariant subset of these elements which forms a basis for $\Lambda^n(A)$. This proves lemma 2.4.

Lemma 2.5. *The exterior powers of indecomposable modules are given by*

$$\begin{aligned}
\Lambda^0(\mathbf{Z}_p) &= \mathbf{Z}_p \\
\Lambda^1(\mathbf{Z}_p) &= \mathbf{Z}_p \\
\Lambda^n(\mathbf{Z}_p) &= 0 \textit{ if } n \neq 0, 1 \\
\Lambda^0(\mathbf{Z}_p[G]) &= \mathbf{Z}_p \\
\Lambda^p(\mathbf{Z}_p[G]) &= \mathbf{Z}_p \textit{ if } p \textit{ is odd, and } I \textit{ if } p = 2 \\
\Lambda^n(\mathbf{Z}_p[G]) &= \mathbf{Z}_p[G]^{\binom{p-1}{n}} \textit{ if } 1 \leq n \leq p-1 \\
\Lambda^n(\mathbf{Z}_p[G]) &= 0 \textit{ if } n > p \\
\Lambda^n(I) &= \mathbf{Z}_p \oplus \mathbf{Z}_p[G]^{(\binom{p-1}{n}-1)/p} \textit{ if } n \textit{ is even and } 0 \leq n < p \\
\Lambda^n(I) &= I \oplus \mathbf{Z}_p[G]^{(\binom{p-1}{n}+1-p)/p} \textit{ if } n \textit{ is odd and } 0 \leq n < p \\
\Lambda^n(I) &= 0 \textit{ if } n \geq p
\end{aligned}$$

Proof. For the exterior powers of $\mathbf{Z}_p[G]$ with p odd we apply lemma 2.4 to see that all exterior powers are sums of $\mathbf{Z}_p$ and $\mathbf{Z}_p[G]$, and the exterior powers can then be identified by checking the dimension and the trace of g. The cohomology groups of the exterior powers of I can then be worked out by induction using the exact sequence

$$0 \to \Lambda^n(I) \to \Lambda^n(\mathbf{Z}_p[G]) \to \Lambda^{n-1}(I) \to 0.$$

The cohomology groups determine how many copies of $\mathbf{Z}_p$ and I occur in the decomposition into indecomposable modules, so we can work out the exterior powers of I from this. The remaining cases of lemma 2.5 are trivial, so this proves lemma 2.5.

These exterior powers induce operators on the ring K which we also denote by Λ^n. For example, we can use these operators to define Adams operations ψ^n on the ring K by

$$\sum_{n \in \mathbf{Z}} (-1)^n \Lambda^n([A]) q^n = \exp(-\sum_{n \in \mathbf{Z}} \psi^n([A]) q^n / n).$$

Corollary 2.6. *If p is odd then*

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(\mathbf{Z}_p)])q^n = 1-q$$

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(\mathbf{Z}_p[G])])q^n = 1-q^p$$

$$\sum_{n\in\mathbf{Z}}(-1)^n f([\Lambda^n(I)])q^n = (1+q^p)/(1+q)$$

Proof. This follows immediately from lemma 2.5 and lemma 2.3.

Lemma 2.7. *The symmetric powers of indecomposable modules are given by*

$$S^n(\mathbf{Z}_p) = \mathbf{Z}_p \text{ if } n \geq 0$$

$$S^n(\mathbf{Z}_p[G]) = \mathbf{Z}_p[G]^{\binom{p+n-1}{n}/p} \text{ if } (p,n)=1$$

$$S^n(\mathbf{Z}_p[G]) = \mathbf{Z}_p[G]^{(\binom{p+n-1}{n}-1)/p} \oplus \mathbf{Z}_p \text{ if } p|n$$

$$S^n(I) = \mathbf{Z}_p \oplus \mathbf{Z}_p[G]^{(\binom{p+n-2}{n}-1)/p} \text{ if } p|n$$

$$S^n(I) = I \oplus \mathbf{Z}_p[G]^{(\binom{p+n-2}{n}+1-p)/p} \text{ if } n \equiv 1 \bmod p$$

$$S^n(I) = \mathbf{Z}_p[G]^{(\binom{p+n-2}{n})/p} \text{ if } n \not\equiv 0,1 \bmod p$$

Proof. For symmetric powers of $\mathbf{Z}_p$ this is trivial and for symmetric powers of $\mathbf{Z}_p[G]$ it can be proved in the same way as in lemma 2.5. If $p=2$ then the case of I is easy to do as I is just one dimensional. For the case of I when p is odd we let N_G be the element $1+g+\cdots+g^{p-1} \in \mathbf{Z}_p[G]$ and consider the quotient of the symmetric algebra $S^*(\mathbf{Z}_p[G])$ by the ideal (N_G). This is the universal commutative algebra generated by $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$, so it is $S^*(\mathbf{Z}_p[G]/N_G\mathbf{Z}_p)$ which is isomorphic to $S^*(I)$ because $\mathbf{Z}_p[G]/N_G\mathbf{Z}_p$ is isomorphic to I. Therefore there is an exact sequence

$$0 \to S^{n-1}(\mathbf{Z}_p[G]) \to S^n(\mathbf{Z}_p[G]) \to S^n(I) \to 0.$$

Using the known cohomology groups of the symmetric powers of $\mathbf{Z}_p[G]$ and the exact sequence of cohomology groups of this exact sequence gives the cohomology groups of $S^n(I)$ and hence the number of times that $\mathbf{Z}_p$ and I occur in the decomposition of $S^n(I)$. This proves lemma 2.7.

Corollary 2.8. *The cohomology rings of the symmetric algebras of indecomposable modules are give as follows.*

$\hat{H}^*(S^*(\mathbf{Z}_p))$ is a polynomial algebra with one ordinary generator of degree 1.

$\hat{H}^*(S^*(\mathbf{Z}_p[G]))$ is a polynomial algebra with one ordinary generator of degree p.

$\hat{H}^*(S^*(I))$ is the tensor product of a polynomial algebra with one ordinary generator of degree p and an exterior algebra with one super generator of degree 1 if p is odd. (If p is even then $\hat{H}^*(S^*(I))$ is a polynomial algebra on one super generator of degree 1, but we do not need this case.)

Proof. This follows from the calculation of the symmetric powers in lemma 2.7. The product structure on the symmetric algebras can be worked out using the formula $d(ab) = (da)b \pm a(db)$ relating the cup product with the map d from $\hat{H}^i(C)$ to $\hat{H}^{i+1}(A)$ associated to the exact sequence $0 \to A \to B \to C \to 0$. This proves corollary 2.8.

Lemma 2.9. *Suppose that $A_{-1}=0, A_0,\ldots,A_n, A_{n+1}=0$ is a sequence of modules and we have maps d from A_i to A_{i+1} and d^* from A_{i+1} to A_i for all i such that $d^2=0$, $d^{*2}=0$, and $dd^*+d^*d=k$ for some fixed integer k coprime to p. Then*

$$[A_0]-[A_1]+[A_2]-\cdots+(-1)^n[A_n]=0.$$

Proof. The fact that k is a unit in $\mathbf{Z}_p$ implies that A_i is the direct sum $dd^*(A_i) \oplus d^*d(A_i)$ and that d is an isomorphism from $d^*d(A_i)$ to $dd^*(A_{i+1})$. The lemma follows from this because the sequence splits as the sum of these isomorphisms. (Notice that the relation between the $[A_i]$'s does not follow just from the exactness of d; for example, look at the short exact sequence $0 \to I \to \mathbf{Z}_p[G] \to \mathbf{Z}_p \to 0$. It is easy to check that in this case there is a map d^* with $d^{*2}=0$ and $dd^*+d^*d=p$, so the condition that $(k,p)=1$ is also necessary.) This proves lemma 2.9.

Lemma 2.10. *Suppose that A is a finitely generated free module over $\mathbf{Z}$ or $\mathbf{Z}_p$ with a symmetric self dual form $(,)$. Then this induces a symmetric self dual form (defined in the proof) on any exterior power $\Lambda^n(A)$.*

Proof. We define the bilinear form on $\Lambda^n(A)$ by

$$(a_1 \wedge \cdots \wedge a_n, b_1 \wedge \cdots \wedge b_n) = \sum_{\sigma \in S_n} \chi(\sigma)(a_1, b_{\sigma(1)}) \cdots (a_n, b_{\sigma(n)})$$

where $\chi(\sigma)$ is 1 or -1 depending on whether σ is an even or odd permutation. This obviously defines a symmetric bilinear form on $\Lambda^n(A)$ with values in $\mathbf{Z}$ or $\mathbf{Z}_p$, and we have to check it is unimodular. The determinant of $\Lambda^n(A)$ is equal to $a\det(A)^b$ for some constants a and b depending on n, but not on A. We can work out a by letting A be the lattice with a base of orthogonal elements $a_1, a_2, \ldots$ with norms ± 1, when we see that $\Lambda^n(A)$ also has an orthogonal base of elements $a_{i_1} \wedge a_{i_2} \cdots$ with $i_1 < i_2 < \cdots$ with norms ± 1, so both A and $\Lambda^n(A)$ have discriminant 1 and so $a = 1$. Hence if A is self dual then $\Lambda^n(A)$ has discriminant 1 and must also be self dual. This proves lemma 2.10. (It is easy to check that the constant b is equal to $\dim(\Lambda^n(A))/\dim(A)$ but we do not need this.)

The analogue of this lemma for symmetric powers is false. The proof breaks down because the element $a_{i_1}a_{i_2}\cdots$ with $i_1 \le i_2 \le \cdots$ does not have norm ± 1 if some of the i_j's are equal. In section 5 we will construct a graded algebra $h(A)$, such that the homogeneous components of $h(A)$ have self dual symmetric bilinear forms if A does, and $h(A) \otimes \mathbf{Q}_p$ is a sum of tensor products of symmetric powers of A.

The following corollary is purely for entertainment and is not used elsewhere in the paper.

Corollary 2.11. *There is a 28-dimensional self dual positive definite lattice with no roots, acted on by the orthogonal group $O_8^+(\mathbf{F}_2)$.*

Proof. By lemma 2.10 the lattice $\Lambda^2(E_8)$ is self dual, and is obviously acted on by the automorphism group of the E_8 lattice modulo -1, which is just $O_8^+(\mathbf{F}_2)$. It is easy to check that this lattice has no roots either by calculation or by observing that there is no suitable root lattice acted on by $O_8^+(\mathbf{F}_2)$. This proves lemma 2.11.

The lattice $\Lambda^2(E_8)$ has theta function

$$\sum_{\lambda \in \Lambda^2(E_8)} q^{\lambda^2/2} = 1 + 2240q^{3/2} + 98280q^{4/2} + 1790208q^{5/2} + 19138560q^{6/2} + \cdots.$$

The vectors of norm 3 are just the 2240 vectors generating one of the 1-dimensional sublattices of the form $\Lambda^2(L)$ where L is one of the 1120 sublattices of E_8 isomorphic to A_2.

3. The $\mathbf{Z}[1/2]$-form of the monster Lie algebra.

We find some relations between the dimensions of the Tate cohomology groups $\hat{H}^i(g, V[1/2])$ by studying the Lie algebra cohomology of the positive subalgebra of a $\mathbf{Z}_p$-form of the monster Lie algebra. In the next section we will use these relations to compute the cohomology groups in the case when g has order at least 13. We recall that $V[1/2]$ is the self dual $\mathbf{Z}[1/2]$-form of Frenkel, Lepowsky, and Meurman's monster vertex algebra [FLM] described in [B-R] (but the arguments in this section work for any self dual $\mathbf{Z}_p$-form of the monster vertex algebra).

This section should really be about the integral form of the monster Lie algebra, but this has not been constructed yet because of the lack of a construction of a self dual integral form of the monster vertex algebra. However, as we are only looking at the action of elements of odd order p in the monster and tensoring everything with $\mathbf{Z}_p$, a $\mathbf{Z}[1/2]$-form is just as good as an integral form.

We briefly recall some properties of the monster Lie algebra $\mathbf{m}$ from [B]. This is a $II_{1,1}$-graded Lie algebra acted on by the monster, where $II_{1,1}$ is the 2-dimensional even self dual Lorentzian lattice whose element (m, n) has norm $-2mn$. The piece of degree (m, n) of $\mathbf{m}$ is isomorphic to V_{mn} if $(m, n) \neq (0, 0)$ and to $\mathbf{R}^2$ if $(m, n) = (0, 0)$. The Lie algebra $\mathbf{m}$ is the sum of 3 subalgebras E, F, and H, where the Cartan subalgebra H is the piece of degree $(0, 0)$ of $\mathbf{m}$, E is the sum of all pieces of degree (m, n) of $\mathbf{m}$ for $m > 0$, and F is the sum of all pieces of degree (m, n) of $\mathbf{m}$ for $m < 0$.

The $\mathbf{Z}[1/2]$ form on V induces a self dual $\mathbf{Z}[1/2]$-form and hence a $\mathbf{Z}_p$-form on $\mathbf{m}$. From now on we will write $\mathbf{m}[1/2]$ for this $\mathbf{Z}[1/2]$-form. In the preprint version of this paper I implicitly assumed that the degree (m, n) piece of $\mathbf{m}$ was self dual and isomorphic as a module over the monster to V_{mn}. At the last moment I realized that this is not at all clear. The problem is that although the no-ghost theorem gives an isomorphism of both spaces tensored with $\mathbf{Q}$, and both spaces have $\mathbf{Z}[1/2]$-forms, there is no obvious reason why this isomorphism should map one $\mathbf{Z}[1/2]$ form to the other. (In fact for the fake monster Lie algebra the corresponding isomorphism does not preserve the integral forms.) Fortunately in this paper we only need the following weaker statement:

Assumption. *If $m < p$ then the degree (m, n) piece of $\mathbf{m} \otimes \mathbf{Z}_p$ is self dual under the natural bilinear form and isomorphic to $V_{mn} \otimes \mathbf{Z}_p$ as a $\mathbf{Z}_p$ module acted on by the monster.*

I believe I have a proof of this; if all goes well it will appear in a paper provisionally titled "The fake monster formal group". Until this paper appears all the statements in sections 3 and 4 and the modular moonshine conjectures for elements of type pA ($p \geq 13$) should have this assumption added as a hypothesis.

I do not know if the condition $m < p$ in the assumption above is necessary; if it is not, then this condition (usually added in parentheses) can be missed out of some of the following lemmas. It seems very likely that self duality still holds without it, but I am not so sure about the action of the monster being the same if both m and n are divisible by p.

The monster Lie algebra has a Weyl vector ρ, i.e., a vector in the root lattice $II_{1,1}$ such that $(\rho, \alpha) = -(\alpha, \alpha)/2$ for every simple root α. This follows from [B, theorem 7.2] which states that this simple roots are the vectors $(1, n)$ for $n \geq 0$ or $n = -1$, with multiplicities $c(n)$. Hence the Weyl vector is $\rho = (1, 0)$.

We recall some facts about Hodge theory applied to Lie algebra cohomology. The self dual symmetric bilinear form on the monster Lie algebra identifies E with the dual space of F, and using the Cartan involution which maps F to E we see that we have a symmetric bilinear form on the Lie algebra E (which is the restriction of the contravariant form of the monster Lie algebra to E) and in particular we can identify E with its dual. The Lie algebra cohomology is the cohomology of an operator d which maps each space $\Lambda^i(E)$ to $\Lambda^{i+1}(E)$. The adjoint of this operator is denoted by d^*, and we define the Laplace operator Ω by $\Omega = dd^* + d^*d$. As for any generalized Kac-Moody algebra with a Weyl vector, the action of the Laplace operator $\Omega = dd^* + d^*d$ on the degree $\alpha = (m, n) \in II_{1,1}$ piece of $\Lambda^*(E)$ is given by multiplication by $(\alpha, \alpha + 2\rho)/2 = (m - 1)n$. The Lie algebra cohomology group $H^i(E)$ of E can be identified with the zero eigenspace of Ω on $\Lambda^i(E)$. The Laplace operator restricted to $\Lambda^1(E) = E$ is half the partial Casimir operator Ω_0 defined in [K, section 2.5]; the extra factor of $1/2$ is necessary to make things work well over rings not containing $1/2$, but this is not important in this paper as we work over the ring of p-adic integers for odd p which always contains $1/2$.

Proposition 3.1. *The operators d and d^* act on the $\mathbf{Z}[1/2]$-form $\Lambda^*(E[1/2]) \subset \Lambda^*(E)$ of the exterior algebra (in all degrees (m, n) with $m < p$).*

Proof. The main point is that all homogeneous pieces of the $\mathbf{Z}[1/2]$-form of the exterior algebra are self dual under the contravariant bilinear form. This follows from lemma 2.10 and the fact that each homogeneous component is a finite sum of tensor products of exterior powers of the homogeneous components of $E[1/2]$. The result now follows from the fact that if d is a homomorphism of finite dimensional free $\mathbf{Z}[1/2]$-modules from A to B and A and B have self dual symmetric inner products then the adjoint d^* from B to A is well defined. This proves proposition 3.1.

We let $R(0)$ be the ring of formal Laurent series $K[[r, q]][1/r, 1/q]$ with coefficients in K, and we let $R(1)$ be the space of power series in $R(0)$ only involving monomials $r^m q^n$ with $p|mn$, and we let $R(2)$ be the subring of $R(0)$ of power series only involving powers of r^p and q^p. So we have the inclusions $R(2) \subset R(1) \subset R(0)$ and $R(1)$ is a module over $R(2)$ (but is not a ring itself). (We can also define $R(0)$, $R(1)$, and $R(2)$ for arbitrary generalized Kac-Moody algebras, in particular for the fake monster Lie algebra: $R(0)$ is a completion of the group ring of the root lattice, $R(1)$ is a subspace corresponding to the elements α of the root lattice with $p|(\alpha, \alpha)/2$, and $R(2)$ is the subring corresponding to p times the root lattice.) If A is a $\mathbf{Z}[1/2]$-module then we write $[A]$ for $[A \otimes_{\mathbf{Z}[1/2]} \mathbf{Z}_p]$.

Lemma 3.2. *If $(p, (m - 1)n) = 1$ (and $m < p$) then the piece of degree (m, n) of*

$$[\Lambda^*(E[1/2])] = [\Lambda^0(E[1/2])] - [\Lambda^1(E[1/2])] + [\Lambda^2(E[1/2])] - \cdots$$

vanishes. In other words

$$r^{-1}\Lambda^*(\sum_{m>0,n\in\mathbf{Z}} r^m q^n [V_{mn}])$$

lies in $R(1)$.

Proof. The Laplace operator $\Omega = dd^*+d^*d$ acts as multiplication by $(m-1)n$ on the degree (m,n) piece of $\Lambda^*(E[1/2])$, so the result follows from lemma 2.9 if we let A_i be the degree (m,n) piece of $\Lambda^i(E[1/2])$. This proves lemma 3.2.

Lemma 3.3. *If* A *is a* $\mathbf{Z}_p[G]$*-module with* $\mathrm{Tr}(g|A) = c^+$, $f([A]) = c^-$, *(and* $m < p$*) then*

$$f(\Lambda^*(r^m q^n A)) = (1 - r^m q^n)^{(c^+ + c^-)/2}(1 + r^m q^n)^{(c^+ - c^-)/2}U$$

where U *is a unit of the ring* $R(2)$.

Proof. It is sufficient to prove this for an indecomposable module A because both sides are multiplicative. We just check each of the three possibilities for A using lemma 2.6.

If $A = \mathbf{Z}_p$ then $c^+ = c^- = 1$ and $f(\Lambda^*(r^m q^n [\mathbf{Z}_p])) = 1 - r^m q^n$, so lemma 3.3 is true in this case.

If $A = \mathbf{Z}_p[G]$ then $c^+ = c^- = 0$ and $f(\Lambda^*(r^m q^n [\mathbf{Z}_p[G]])) = 1 - r^{mp}q^{np}$ which is a unit in $R(2)$, so lemma 3.3 is true in this case.

If $A = I$ then $c^+ = -1$, $c^- = 1$ and

$$f(\Lambda^*(r^m q^n [I])) = (1 + r^{mp}q^{np})/(1 + r^m q^n) = U(1 + r^m q^n)^{-1},$$

so lemma 3.3 is also true in this case. This proves lemma 3.3.

We define integers $c_g^+(n)$ and $c_g^-(n)$ by

$$\begin{aligned} c_g^+(n) &= \mathrm{Tr}(g|V_n) = \dim(\hat{H}^0(g, V[1/2]_n)) - \dim(\hat{H}^1(g, V[1/2]_n)) \\ c_g^-(n) &= f(V[1/2]_n) = \dim(\hat{H}^0(g, V[1/2]_n)) + \dim(\hat{H}^1(g, V[1/2]_n)). \end{aligned}$$

We use the numbers $c_g^+(n)$ and $c_g^-(n)$ rather than the apparently simpler numbers $\dim(\hat{H}^0(g, V[1/2]_n))$ and $\dim(\hat{H}^1(g, V[1/2]_n))$ because they turn out to be coefficients of Hauptmoduls, and because $c_g^+(n)$ is already known.

Proposition 3.4. *If* $(p, mn) = 1$ *(and* $m < p$*) then the coefficient of* $r^m q^n$ *in*

$$r^{-1}\prod_{m>0,n\in\mathbf{Z}} (1 - r^m q^n)^{(c_g^-(mn)+c_g^+(mn))/2}(1 + r^m q^n)^{(c_g^+(mn)-c_g^-(mn))/2}$$

vanishes. In other words this power series lies in $R(1)$.

Proof. We apply the homomorphism f to the expression in lemma 3.2 and use lemma 3.3. We find that the expression in this proposition is equal to an element of $R(1)$ times a unit in $R(2)$, and is therefore still an element of $R(1)$. Therefore its coefficients of $r^m q^n$ vanish unless $p|mn$. This proves proposition 3.4.

Proposition 3.4 can be generalized in the obvious way to any generalized Kac-Moody algebra which has a self dual $\mathbf{Z}_p$-form, a Weyl vector, and an integral root lattice.

We can also ask whether or not the coefficients of $r^m q^n$ in proposition 3.4 vanish when $p|mn$. Some numerical calculations suggest that they usually do not.

4. The modular moonshine conjectures for $p \geq 13$.

In this section we prove the following theorem, which completes the proof of the modular moonshine conjectures of [R section 6] (apart from a small technicality in the case $p = 2$.)

Theorem 4.1. *If $g \in M$ is an element of prime order $p \geq 17$ or an element of type $13A$ then $\hat{H}^1(g, V[1/2]) = 0$.*

This implies that ${}^gV = \hat{H}^0(g, V[1/2]) = \hat{H}^*(g, V[1/2])$ is a vertex algebra whose modular character is given by Hauptmoduls, and whose homogeneous components have the characters of [R definition 2].

The proof of this theorem will occupy the rest of this section. We have to show that the numbers $c_g^-(n)$ of proposition 3.4 are equal to the numbers $c_g^+(n)$, because the difference is twice the dimension of $\hat{H}^1(g, V[1/2]_n)$. We start by summarizing what we know about these numbers.

Lemma 4.2.

1. *The numbers $c_g^-(n)$ and $c_g^+(n)$ satisfy the relations given in proposition 3.4.*
2. *The numbers $c_g^-(n)$ are integers, with $c_g^-(n) \equiv c_g^+(n) \bmod 2$.*
3. *$c_g^-(n) \geq |c_g^+(n)|$ and $(p-2)c_g^+(n) + pc_g^-(n) \leq 2c_1^+(n)$, (where $c_1^+(n)$ is the coefficient of q^n in the elliptic modular function).*
4. *The numbers $c_g^+(n)$ are the coefficients of the Hauptmodul of the element $g \in M$.*

Proof. These properties follow easily from proposition 3.4 together with the fact that $\mathbf{Z}_p \otimes V_n$ has dimension $c_1^+(n)$ and is the sum of $(c_g^-(n) + c_g^+(n))/2$ copies of $\mathbf{Z}_p$, $(c_g^-(n) - c_g^+(n))/2$ copies of I, and some copies of $\mathbf{Z}_p[G]$. This proves lemma 4.2.

We will prove theorem 4.1 by showing that if g satisfies the conditions of theorem 4.1 then the conditions 1 to 4 in lemma 4.2 imply that $c_g^-(n) = c_g^+(n)$. We find a finite set of possible solutions of conditions 1, 2, and 4 of lemma 4.2, and find that for each p only one of these satisfies the condition 3. (There are often several solutions not satisfying condition 3.) The proof is just a long messy calculation and the reader should not waste time looking at it.

We first put the relations into a more convenient form.

Proposition 4.3. *There are integers $c_{m,n}$ defined for $m, n > 0$ (and $m < p$) with the following properties.*

1. $$-\log\left(1 + \sum_{m>0,n>0} c_{m,n} r^m q^n\right) \equiv \sum_{m>0,n>0} \sum_{d|(m,n)} \frac{c_g^{(-)^d}(mn/d^2)}{d} r^m q^n \bmod r^p$$
2. *If $(p, mn) = 1$ then $c_{m+1,n} = c_{m,n+1}$.*

Proof. This follows from proposition 3.4 if we multiply both sides by $(r-q)/rq$ and then take the logarithm of both sides. This proves proposition 4.3.

Lemma 4.4. *If $p > 13$ then the values of $c_g^-(n)$ for $n \leq 21$, $c_g^-(2n)$ for $2n \leq 32$, and $c_g^-(36)$ and $c_g^-(45)$ are given by polynomials in the numbers $c_g^-(i)$ for $i = 1, 2, 4, 5$ with coefficients in $\mathbf{Z}_3$.*

Proof. We can evaluate the elements $c(n), n \leq 21$ and $c(2n), n \leq 16$ using the argument for cases 1 and 2 of lemma 4.7 below. We can evaluate $c(36)$ by looking at the coefficient of r^4q^9 in proposition 4.3 and using the fact that we know $c_{m,n}$ for $m \leq 3$, $n \leq 8$. Similarly we can evaluate $c(45)$ by looking at the coefficient of r^5q^9. Notice that the term $(\sum_{m,n} c_{m,n} r^m q^n)^3/3$ has 3's only in the denominators of coefficients of $r^m q^n$ when $3|m$ and $3|n$, so we do not get problems from this for the coefficients of r^4q^9 and r^5q^9 (but we do get problems from this if we try to work out $c_g^-(27)$ using the same method). This proves lemma 4.4.

If $p = 13$ we run into trouble when we try to determine $c_g^-(26)$; this is why the argument in lemma 4.4 does not work in this case. Also the argument breaks down if we try to work out $c(27)$, because when we look at the coefficient of r^3q^9 we get an extra term $c_g^-(3)/3$ which does not have coefficients in $\mathbf{Z}_3$. This is why we use the coefficients $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ rather than $c_g^-(1)$, $c_g^-(2)$, $c_g^-(3)$, and $c_g^-(5)$.

Lemma 4.4 is the reason why our arguments do not work for $p < 13$ (and do not work so smoothly for $p = 13$). As p gets smaller we have fewer relations to work with and it gets more difficult to determine all coefficients in terms of the $c_g^-(m)$'s for small m. It is probably possible to extend lemma 4.4 to cover some smaller primes with a lot of effort. Fortunately it is not necessary to do the cases $p < 13$ because these cases have already been done by explicit calculations in [B-R].

Lemma 4.5. *If $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are known mod 3^n for some $n \geq 1$, and $p > 13$, and $(c_g^-(1), 3) = 1$ if $p = 13$, then these values of $c_g^-(m)$ are determined mod 3^{n+1}.*

Proof. By lemma 4.4 we see that we can determine the values mod 3^n of $c_g^-(n)$ for $n \leq 16$, $c_g^-(2n)$ for $2n \leq 32$, and $c_g^-(36)$ and $c_g^-(45)$.

But now if we look at the coefficient of r^3q^{3m} of 3.4 with $m = 1, 2, 4$, or 5, we see that $c_g^-(9m) + c_g^+(2m)/2 + c_g^-(m)/3 = c_g^-(m)^3/3+$ (some polynomial in known $c_g^-(i)$'s with coefficients that are 3-adic integers). But $c_g^-(m)^3/3$ mod 3^n depends only on $c_g^-(m)$ mod 3^n if $n \geq 1$, so we can determine $c_g^-(m)/3$ mod 3^n and hence $c_g^-(m)$ mod 3^{n+1}. This proves lemma 4.5.

This lemma is the reason that we use 3-adic rather than 2-adic approximation. If we try to prove the lemma above for 2^{n+1} instead of 3^{n+1}, all we find is that we can determine the $c_g^+(m)$'s mod 2^{n+1}, which is useless because we already know these numbers.

Proposition 4.6. *If the numbers $c_g^-(n)$ satisfy the conditions 1, 2, and 4 of lemma 4.2 then the numbers $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are congruent mod 3^{29} to the coefficients of q, q^2, q^4, and q^5 of one of the following power series. (The second column is a genus zero group whose Hauptmodul appears to be the function with coefficients $c_g^-(n)$. This has not been checked rigorously as it is not necessary for the proof of theorem 4.1.)*

$p = 71$	$\Gamma_0(71)+$	q^{-1}	$+q$	$+q^2$	$+q^3$	$+q^4$	$+2q^5$	$+\cdots$
$p = 59$	$\Gamma_0(59)+$	q^{-1}	$+q$	$+q^2$	$+2q^3$	$+2q^4$	$+3q^5$	$+\cdots$
$p = 47$	$\Gamma_0(47)+$	q^{-1}	$+q$	$+2q^2$	$+3q^3$	$+3q^4$	$+5q^5$	$+\cdots$
$p = 47$	$\Gamma_0(94)+$	q^{-1}	$+q$		$+q^3$	$+q^4$	$+q^5$	$+\cdots$
$p = 41$	$\Gamma_0(41)+$	q^{-1}	$+2q$	$+2q^2$	$+3q^3$	$+4q^4$	$+7q^5$	$+\cdots$
$p = 41$	$\Gamma_0(82\|2)+$	q^{-1}			$+q^3$		$+q^5$	$+\cdots$
$p = 31$	$\Gamma_0(31)+$	q^{-1}	$+3q$	$+3q^2$	$+6q^3$	$+9q^4$	$+13q^5$	$+\cdots$
$p = 31$	$\Gamma_0(62)+$	q^{-1}	$+q$	$+q^2$	$+2q^3$	$+q^4$	$+3q^5$	$+\cdots$
$p = 29$	$\Gamma_0(29)+$	q^{-1}	$+3q$	$+4q^2$	$+7q^3$	$+10q^4$	$+17q^5$	$+\cdots$
$p = 29$	$\Gamma_0(58\|2)+$	q^{-1}	$+q$		$+q^3$		$+q^5$	$+\cdots$
$p = 23$	$\Gamma_0(23)+$	q^{-1}	$+4q$	$+7q^2$	$+13q^3$	$+19q^4$	$+33q^5$	$+\cdots$
$p = 23$	$\Gamma_0(46) + 23$	q^{-1}		$-q^2$	$+q^3$	$-q^4$	$+q^5$	$+\cdots$
$p = 23$	$\Gamma_0(46)+$	q^{-1}	$+2q$	$+q^2$	$+3q^3$	$+3q^4$	$+5q^5$	$+\cdots$
$p = 19$	$\Gamma_0(19)+$	q^{-1}	$+6q$	$+10q^2$	$+21q^3$	$+36q^4$	$+61q^5$	$+\cdots$
$p = 19$	$\Gamma_0(38)+$	q^{-1}	$+2q$	$+2q^2$	$+5q^3$	$+4q^4$	$+9q^5$	$+\cdots$
$p = 19$	$\Gamma_0(38\|2)+$	q^{-1}	$+2q$		$+q^3$		$+3q^5$	$+\cdots$
$p = 17$	$\Gamma_0(17)+$	q^{-1}	$+7q$	$+14q^2$	$+29q^3$	$+50q^4$	$+92q^5$	$+\cdots$
$p = 17$	$\Gamma_0(34)+$	q^{-1}	$+3q$	$+2q^2$	$+5q^3$	$+6q^4$	$+12q^5$	$+\cdots$
$p = 17$	$\Gamma_0(34\|2)+$	q^{-1}	$+q$		$+3q^3$		$+4q^5$	$+\cdots$
$p = 13$	$\Gamma_0(13)+$	q^{-1}	$+12q$	$+28q^2$	$+66q^3$	$+132q^4$	$+258q^5$	$+\cdots$
$p = 13$	$\Gamma_0(26)+$	q^{-1}	$+4q$	$+4q^2$	$+10q^3$	$+12q^4$	$+26q^5$	$+\cdots$
$p = 13$	$\Gamma_0(26\|2)+$	q^{-1}	$+2q$		$+4q^3$		$+6q^5$	$+\cdots$

It is also possible (but unlikely) that there are other solutions for $p = 13$ for which $c_g^-(1)$ does not satisfy the inequalities $0 \leq c_g^-(1) < 3^{10}$.

Proof. For each prime p we use a computer to test all 81 possibilities for $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ mod 3. If $p > 13$ then we can calculate the p-adic expansion of all the coefficients $c_g^-(n)$ recursively using lemma 4.5 above, and we reach a contradiction by looking at coefficients of proposition 4.3 except in the cases above. For $p = 13$ this does not quite work as we have not shown the values mod 3^n determine those mod 3^{n+1}, so we can adopt the crude procedure of just testing all 3^4 possibilities for the $c_g^-(i)$'s mod 3^{n+1} for each solution mod 3^n we have found, and checking to see which of them leads to contradictions. There were some cases for $p = 13$ where this did not lead to a contradiction but did at least lead to the conclusion that $c_g^-(1) < 0$ or $c_g^-(1) \geq 3^{10}$. (If $c_g^-(1)$ is not divisible by 3 then even when $p = 13$ the numbers $c_g^-(i)$ mod 3^n determine the numbers $c_g^-(n)$ mod 3^{n+1}. When $c_g^-(1)$ was divisible by 3, there were several cases where the coefficient $c_g^-(5)$ was not determined mod 3^{n+1} by the identities used by the computer program. However in all the cases looked at by the computer with $3|c_g^-(1)$, the values of $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $3c_g^-(5)$ mod 3^n uniquely determined their values mod 3^{n+1}.) This proves proposition 4.6, at least if one believes the computer calculations. (Anyone who does not like computer calculations in a proof is welcome to redo the calculations by hand.)

Lemma 4.7. *If the coefficients $c_g^-(1)$, $c_g^-(2)$, $c_g^-(4)$, and $c_g^-(5)$ are equal to the coefficients $c_g^+(1)$, $c_g^+(2)$, $c_g^+(4)$, and $c_g^+(5)$ for g an element of prime order $p > 13$ or an element of type $13A$ then all the coefficients are determined by the relations of proposition 3.4.*

Proof. This proof is just a long case by case check using induction. We will repeatedly use the fact that the coefficients of $r^m q^n$ of both sides of proposition 4.3 are equal.

Case 1. $c_g^-(2n)$ for $2n \not\equiv 0 \bmod p$. By looking at the coefficient of r^2q^n and using the fact that $c_{2,n} = c_{1,n+1}$ (as $(p,n) = 1$) we see that $c_g^-(2n)$ can be written as a polynomial in $c(n+1)$ and $c(i)$ for $1 \leq i < n$.

Case 2. $c_g^-(2n+1)$ for $2n+1 \not\equiv \pm 1 \bmod p$. By looking at the coefficient of r^2q^{2n} we can express $c_g^-(2n+1)$ in terms of $c_g^-(4n)$ and known quantities. But $c_g^-(4n)$ can be evaluated by looking at the coefficient of r^4q^n and using the fact that $c_{4,n} = c_{3,n+1} = c_{2,n+2}$ (as $2n+1 \not\equiv \pm 1 \bmod p$).

The recursion formulas given by cases 1 or 2 can be written out explicitly and are given in [B, (9.1)]. In that paper they are sufficient to determine all coefficients because the analogues of the relations in proposition 3.4 hold whenever $mn \neq 0$.

The remaining cases of the proof are unusually tedious so we only sketch them briefly.

Case 3. $c_g^-(2n)$ for $2n \equiv 0 \bmod p$, n odd. We can express this in terms of $c_{2,n} = c_{3,n-1}$, which we can evaluate by calculating $c_g^-(3(n-1)) = c_g^-(2(3(n-1)/2))$ using case 1.

For the rest of the proof it is convenient to evaluate the cluster of elements $c_g^-(2pn-1)$, $c_g^-(2pn+1)$, and $c_g^-(4pn)$, simultaneously. (These are the only cases left to do.)

Case 4. $2np \equiv 2 \bmod 3$. We evaluate $c_g^-(2np+1)$ in terms of $c_{3,(2np+1)/3}$. We evaluate $c_g^-(2np-1)$ by expression $c_g^-(2np-1)c(2)$ in terms of $c_{2,2np+1}$ and evaluating the latter by comparing it with $c_{3,2(2np+1)/3}$. We evaluate $c_g^-(4np)$ by expressing $c_{3,2np+2}$ in terms of $c_g^-(2)c_{2,2np}$ and evaluating the latter by comparing it with $c_{2,3np+3}$.

Case 5. $2np \equiv 1 \bmod 3$. We evaluate $c_g^-(2np-1)$ in terms of $c_{3,(2np-1)/3}$. We evaluate $c_{1,2np+5}$ by using $c_{3,(2np+5)/3}$ and then evaluate in turn $c_{2,2np+4} = c_{1,2np+5}$, $c_g^-(1)c_{1,2np+3}$ (using the coefficient of r^2q^{2np+4}), $c_{2,2np+3} = c_{1,2np+3}$, $c_g^-(1)c_{1,2np+1}$ which gives us the value of $c_{1,2np+1}$. We evaluate $c_{2,2np}$ as in case 4.

Case 6. $2np \equiv 0 \bmod 3$. By looking at $c_{2,2np+2}$ and $c_{2,2np+3}$ we can find expressions for the linear combinations $c_g^-(1)c_g^-(2np+1) + c_g^-(3)c_g^-(2np-1)$ and $c_g^-(2)c_g^-(2np+1) + c_g^-(4)c_g^-(2np-1)$. If $p \neq 59$ or 71 then these two linear equations are independent so we can solve for $c_g^-(2np+1)$ and $c_g^-(2np-1)$, and then find $c_{2,2np}$ as in case 4. If $p = 59$ or 71 then an even more tedious argument using $c_{2,2np+2}$, $c_{3,2np+2}$, and $c_{3,2np+3}$ produces 3 independent linear relations between $c_g^-(2pn-1)$, $c_g^-(2pn+1)$, and $c_g^-(4pn)$, so we can solve for them.

This proves lemma 4.7.

We can now complete the proof of theorem 4.1. If we look at the possible solutions for the coefficients $c_g^-(n)$ given by proposition 4.6, we see that all the solutions except those congruent mod 3^{29} to the Hauptmoduls of the elements $g \in M$ are ruled out by the inequalities of lemma 4.2. As $3^{29} > c_1^+(i) > c_g^-(i) \geq 0$ for $i \leq 5$ we see that $c_g^+(i) = c_g^-(i)$ for $i \leq 5$. The coefficients $c_g^+(n)$ also satisfy the identities of proposition 4.3 satisfied by the $c_g^-(n)$'s; for example, this follows easily from the identity

$$\sum_{n\in\mathbf{Z}} c_g^+(n)r^n - \sum_{n\in\mathbf{Z}} c_g^+(n)q^n = r^{-1} \prod_{m>0, n\in\mathbf{Z}} (1 - r^m q^n)^{c_g^+(mn)} (1 - r^{pm}q^{pn})^{c_g^+(pmn)}$$

[B, page 434]. Therefore $c_g^-(n) = c_g^+(n)$ for all n by lemma 4.7 because they both have the same values for $n \leq 5$ and both satisfy the same recursion relations given in proposition 3.4. This proves theorem 4.1.

5. Calculation of some cohomology groups.

In this section we calculate the Tate cohomology groups of some spaces that appear in the construction of the monster vertex algebra. In the next section we will use these calculations to find the Tate cohomology $\hat{H}^*(g, V[1/2])$ for elements g of types $3B$, $5B$, $7B$ and $13B$.

We will write $g_0, g_1, \dots, g_{p-1}$ for the elements $1, g, \dots, g^{p-1}$ of the group ring $\mathbf{Z}_p[G]$ to avoid confusing the group ring multiplication with the multiplication in other rings containing it.

Suppose that A is a free $\mathbf{Z}_p$-module. We write $h(A)$ for the $\mathbf{Z}$-graded $\mathbf{Z}_p$-algebra generated by elements $h_n(a)$ of degree n for $a \in A$, $n \geq 0$ with the relations

1. $h_n(\lambda a) = \lambda^n h_n(a)$ if $\lambda \in \mathbf{Z}_p$.
2. $h_0(a) = 1$, $h_n(a) = 0$ if $n < 0$.
3. $h_n(a+b) = \sum_{i \in \mathbf{Z}} h_i(a) h_{n-i}(b)$.

If A has a basis $a_1, \dots, a_m$ then $h(A)$ is the polynomial algebra generated by elements $h_n(a_i)$ for $n > 0$, $1 \leq i \leq m$. An action of G on A induces an action of G on $h(A)$ in the obvious way.

Lemma 5.1. *The cohomology ring $\hat{H}^*(g, h(\mathbf{Z}_p[G]))$ is the polynomial algebra over $\mathbf{F}_p$ generated by the ordinary elements $h_n(g_0) + h_n(g_1) + \cdots + h_n(g_{p-1})$ for $n > 0$.*

Proof. This follows from the fact that $h(\mathbf{Z}_p[G])$ is the tensor product of the rings A_n, where A_n is the polynomial algebra generated by the elements $h_n(g_i)$ for $0 \leq i < p$. We see that A_n is isomorphic to the symmetric algebra $S^*(\mathbf{Z}_p[G])$ whose cohomology is given by lemma 2.8, and we then apply corollary 2.2 to finish the proof. (Notice that the space spanned by the elements $h_n(g_i)$ for $0 \leq i < p$ is not $h_n(\mathbf{Z}_p[G])$; in fact, the latter is not even a $\mathbf{Z}_p$-module.) This proves lemma 5.1.

Proposition 5.2. *The cohomology ring $\hat{H}^*(g, h(I))$ is isomorphic to the free exterior algebra with one super generator in each positive degree not divisible by p. The automorphism -1 of I multiplies each generator by -1 and hence acts as -1 on $\hat{H}^1(g, h(I))$ and as 1 on $\hat{H}^0(g, h(I))$.*

Proof. The algebra $h(I)$ is the quotient of $h(\mathbf{Z}_p[G])$ by the ideal generated by the elements $h_n(N_G)$. We let R_m be the quotient of $h(\mathbf{Z}_p[G])$ by the elements $h_n(N_G)$ for $n \leq m$, and we will prove by induction on m that $\hat{H}^*(R_m)$ is the tensor product of the free polynomial algebra generated by the ordinary elements $h_n(g_0) + h_n(g_1) + \cdots + h_n(g_{p-1})$ (of degree n) for $n > m/p$ and the free exterior algebra generated by super elements of each degree n with $0 < n \leq m$, $(p, n) = 1$ (of degree n).

To do this we look at the exact sequence

$$0 \to R_{m-1} \to R_{m-1} \to R_m \to 0$$

where the map from R_{m-1} to R_{m-1} is given by multiplication by $h_m(N_G)$, and work out the cohomology of R_m using the cohomology exact sequence of this short exact sequence. We have to consider the cases $p|m$ and $(p, m) = 1$ separately.

If $(p, m) = 1$ then multiplication by $h_m(N_G)$ induces the zero map on cohomology, so that $\hat{H}^*(g, R_m)$ is the sum of $\hat{H}^*(g, R_{m-1})$ and $\hat{H}^*(g, R_{m-1})$ with the degree shifted by m and with the ordinary and super parts exchanged. In other words, $\hat{H}^*(g, R_m)$ is the tensor product of $\hat{H}^*(g, R_{m-1})$ with an exterior algebra generated by one super element of degree m. This proves the induction step when $(p, m) = 1$.

If $p|m$ then multiplication by $h_m(N_G)$ induces an injective map on cohomology, so that $\hat{H}^*(g, R_m)$ is the quotient of $\hat{H}^*(g, R_{m-1})$ by the image of $h_m(N_G)$, in other words it is just $\hat{H}^*(g, R_{m-1})$ with a polynomial generator of degree m killed off.

Hence we see that to go from $\hat{H}^*(h(R_{m-1}))$ to $\hat{H}^*(h(R_m))$ we either kill off a polynomial generator or add an exterior algebra generator. This proves the statement about $\hat{H}^*(h(R_m))$ for all m by induction, and hence proves proposition 5.2.

For completeness we remark that the cohomology ring $\hat{H}^*(g, h(\mathbf{Z}_p))$ is a polynomial algebra over $\mathbf{F}_p$ with one ordinary generator of each positive degree.

For any $\mathbf{Z}_p$-module A we define ω to be the automorphism of $h(A)$ taking $h_n(a)$ to $(-1)^n h_n(-a)$. We define $h_\omega(A)$ to be the largest quotient ring of $h(A)$ on which the automorphism ω acts trivially. (Notice that this is smaller than the largest quotient module on which ω acts trivially.)

Lemma 5.3. *The cohomology ring $\hat{H}^*(g, h_\omega(\mathbf{Z}_p[G]))$ is the polynomial ring on ordinary generators of degrees $(2n+1)p$ for $n \geq 0$.*

Proof. The algebra $h_\omega(\mathbf{Z}_p[G])$ is the polynomial algebra generated by the elements $h_{2n+1}(g_i)$ for $0 \leq i < p$, $n \geq 0$, and lemma 5.3 follows from this.

Lemma 5.4. *The cohomology ring $\hat{H}^*(g, h_\omega(I))$ is the exterior algebra on generators of degrees $(2n+1)$ for $n \geq 0$, $(n,p) = 1$. In particular it is generated (as a ring) by super elements of odd degrees, so that $\hat{H}^0(g, h_\omega(I))$ vanishes in odd degrees and $\hat{H}^1(g, h_\omega(I))$ vanishes in even degrees.*

This follows from lemma 5.3 in the same way that lemma 5.2 follows from 5.1.

Remark. The ring $h(\mathbf{Z}_p)$ can be identified with the ring of symmetric functions over $\mathbf{Z}_p$ if we identify $h_n(1)$ with the n'th complete symmetric function. In this case the automorphism ω taking $h_n(1)$ to $(-1)^n h_n(-1)$ is just the automorphism ω of [Mac, p. 14] taking the complete symmetric functions to the elementary symmetric functions.

Remark. More generally we can form a quotient ring associated to any automorphism ω of A of order n prime to p such that $h_\omega(A)$ is the quotient ring associated to the automorphism $\omega = -1$ of A. To do this we let R_p be the unramified extension of $\mathbf{Z}_p$ generated by a primitive n'th root of unity ζ. We then form the largest quotient ring of $h(A) \otimes R_p$ whose elements are fixed by the automorphism taking $h_n(a)$ to $\zeta^n h_n(\omega(a))$. For example this space appears when we try to construct the monster vertex algebra from a fixed point free automorphism ω of the Leech lattice Λ.

The extraspecial group 2^{1+24} occurring in the centralizer of an element of type $2B$ in M has a unique 2^{12}-dimensional irreducible module A over $\mathbf{Z}[1/2]$. Any element of odd order in $\mathrm{Aut}(\Lambda)$ has an induced action on 2^{1+24} and on A.

Lemma 5.5. *If g is an element of order 3, 5, 7, or 13 acting on the Leech lattice Λ with no nonzero fixed vectors then as a $\mathbf{Z}_p[G]$-module, $A \otimes \mathbf{Z}_p$ is the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$.*

Proof. We let B be the 2^{24}-dimensional 2^{1+24}-module $A \otimes A$. This is isomorphic to the group ring of 2^{24}. The element g acts fixed point freely on the group $2^{24} = \Lambda/2\Lambda$ because it acts fixed point freely on Λ and has odd order, so as a module over G, $B \otimes \mathbf{Z}_p$ is the sum of $\mathbf{Z}_p$ and some copies of $\mathbf{Z}_p[G]$, and the Tate cohomology groups are therefore $\hat{H}^0(g, B \otimes \mathbf{Z}_p) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g, B \otimes \mathbf{Z}_p) = 0$. As $\hat{H}^*(g, B \otimes \mathbf{Z}_p) = \hat{H}^*(g, A \otimes \mathbf{Z}_p) \otimes \hat{H}^*(g, A \otimes \mathbf{Z}_p)$ we see that $\hat{H}^*(g, A \otimes \mathbf{Z}_p)$ is either an ordinary or a super vector space over $\mathbf{F}_p$ of dimension 1.

Therefore as a G module, $A \otimes \mathbf{Z}_p$ is the sum of some copies of $\mathbf{Z}_p[G]$ and either $\mathbf{Z}_p$ or I. But $\dim(A) = 2^{12} \equiv 1 \bmod p$ as $(p-1)|12$ which rules out the second possibility. Therefore $\hat{H}^0(g, A \otimes \mathbf{Z}_p) = \mathbf{Z}/p\mathbf{Z}$, $\hat{H}^1(g, A \otimes \mathbf{Z}_p) = 0$. This proves lemma 5.5.

A similar argument can be used to calculate $\hat{H}^*(g, A)$ for any element g of odd prime order p in $\mathrm{Aut}(\Lambda)$. If $2n$ is the dimension of the subspace of Λ fixed by g then n is an integer with $2^n \equiv \pm 1 \bmod p$. Exactly one of the two cohomology groups $\hat{H}^*(g, A)$ vanishes and the other has dimension 2^n over $\mathbf{F}_p$. The group $\hat{H}^1(g, A)$ vanishes if $2^n \equiv 1 \bmod p$, and the group $\hat{H}^0(g, A)$ vanishes if $2^n \equiv -1 \bmod p$. There are some cases where it is $\hat{H}^0(g, A)$ and not $\hat{H}^1(g, A)$ that vanishes; for example, if g is the element of order 5 with trace -1 on Λ then $\hat{H}^0(g, A) = 0$ and $\hat{H}^1(g, A) = (\mathbf{Z}/5\mathbf{Z})^4$.

6. Elements of type $3B$, $5B$, $7B$, and $13B$.

We use the result of the previous section to calculate the cohomology groups $\hat{H}^i(g, V[1/2])$ when $g \in M$ has type $3B$, $5B$, $7B$, or $13B$.

Theorem 6.1. *If $g \in M$ has type $3B$, $5B$, $7B$, or $13B$ and σ is the element of order 2 generating the center of $C_M(g)/O_p(C_M(g))$ then $O_p(C_M(g))$ acts trivially on ${}^gV$ and σ fixes the ordinary part of ${}^gV$ and acts as -1 on the super part.*

Proof. We prove this by calculating the action of σ on $\hat{H}^*(g, V[1/2])$ explicitly. We recall from [B-R] that the $\mathbf{Z}[1/2]$-form $V[1/2]$ of Frenkel, Lepowsky and Meurman's monster vertex algebra [FLM] is the sum of two pieces V^0 and V^1 which are the $+1$ and -1 eigenspaces of an element σ of type $2B$. (Unfortunately the details for the construction of this $\mathbf{Z}[1/2]$-form have not been published anywhere.) We now show that $\hat{H}^0(g, V^1) = \hat{H}^1(g, V^0) = 0$.

The space V^0 is the subspace of $V_\Lambda[1/2]$ fixed by the automorphism induced by -1 acting on the Leech lattice Λ, where $V_\Lambda[1/2] = V_\Lambda \otimes \mathbf{Z}[1/2]$, and V_Λ is the integral form of the vertex algebra of the Leech lattice, described in [B-R]. As in [B-R], the space V_Λ is the tensor product of a twisted group ring $\mathbf{Z}[\hat{\Lambda}]$ and a space isomorphic to $h(\Lambda)$. The cohomology $\hat{H}^*(g, \mathbf{Z}[\hat{\Lambda}])$ is just $\mathbf{Z}/p\mathbf{Z}$ as in [B-R] because g acts fixed point freely on Λ. By proposition 5.2, the element -1 of $\mathrm{Aut}(\hat\Lambda)$ acts as $+1$ on $\hat H^0(g,h(\Lambda))$ and as -1 on $\hat H^1(g,h(\Lambda))$. Therefore $\hat H^1(g,V^0)$ vanishes because it is isomorphic to the subspace of $\hat H^1(g,h(\Lambda))$ fixed by $-1\in\mathrm{Aut}(\Lambda)$.

The space V^1 is the subspace of elements of integral weight of the tensor product $A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p)$, where A is the 2^{12} dimensional module for 2^{1+24} and $h_\omega(\Lambda\otimes\mathbf{Z}_p)$ has its grading halved and then shifted by $3/2$. By lemma 5.5, $\hat H^*(g,A)=\mathbf{F}_p$, so $\hat H^*(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))=\hat H^*(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$. By lemma 5.4 the cohomology ring $\hat H^*(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ is generated by super elements of odd degree, so $\hat H^1(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in even degrees and $\hat H^1(g,h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in odd degrees. Therefore $\hat H^0(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))$ vanishes in integral degrees, so $\hat H^0(g,V^1)=0$.

This shows that the element σ acts as 1 on $\hat H^0(g,V[1/2])=\hat H^0(g,V_\Lambda)=\hat H^0(g,h(\Lambda\otimes\mathbf{Z}_p))$ and as -1 on $\hat H^1(g,V[1/2])=\hat H^1(g,A\otimes h_\omega(\Lambda\otimes\mathbf{Z}_p))$.

Finally we show that $O_p(C_M(g))$ acts trivially on ${}^gV$. The element σ acts trivially on the center $\langle g\rangle$ of the extraspecial group $p^{1+24/(p-1)}$ and acts as -1 on the quotient $p^{24/(p-1)}$. The element g acts trivially on ${}^gV$ so we get an action of $p^{24/(p-1)}$ on ${}^gV$. As σ acts as 1 on the even part of ${}^gV$ and as -1 on the odd part, it commutes with all automorphisms of ${}^gV$, so the group $p^{24/(p-1)}$ must act trivially on ${}^gV$. This proves theorem 6.1.

The group $C_M(g)$ has structure $3^{1+12}.2.Suz$, $5^{1+6}.2.HJ$, $7^{1+4}.2.A_7$, or $13^{1+2}.2.A_4$, so we get modular vertex superalgebras acted on faithfully by the groups $2.Suz$, $2.HJ$, $2.A_7$, and $2.A_4$ whose modular characters are given by Hauptmoduls. We can calculate the characters of both the ordinary and super parts of these superalgebras by using the fact that the characters of their sum and difference are both given in terms of known Hauptmoduls.

It seems likely that if g has type $2B$ then the vertex superalgebra acted on by $2^{24}.Co_1$ constructed in [B-R] is acted on trivially by the 2^{24} subgroup, but this has not yet been proved. The proof given above for the corresponding fact for elements of odd prime order fails for order 2 because an automorphism mapping each element of 2^{24} to its inverse is still the identity automorphism.

As an example, we work out the dimensions of the ordinary and super parts of ${}^gV$ in the case when g has type $3B$. The function $\sum_n \mathrm{Tr}(1|{}^gV)$ is equal to the Hauptmodul of the element g of type $3B$ in the monster, which is

$$q^{-1}+54q-76q^2-243q^3+1188q^4-1384q^5-\cdots,$$

and the function $\sum_n \mathrm{Tr}(\sigma|{}^gV)$ is equal to the Hauptmodul of the element σg of type $6B$ in the monster, which is

$$q^{-1}+78q+364q^2+1365q^3+4380q^4+12520q^5+\cdots$$

so we find that the dimension of the ordinary part of ${}^gV_n$ is the coefficient of q^n of

$$q^{-1}+66q+144q^2+561q^3+2784q^4+5568q^5+\cdots$$

and the dimension of the super part of ${}^gV_n$ is the coefficient of q^n of

$$12q+220q^2+804q^3+1596q^4+6952q^5\cdots.$$

These series can be written as sums of dimensions of ordinary representations of the groups $3.Suz$ and $6.Suz$ as

$$q^{-1}+66q+(1+143)q^2+(429+66+66)q^3+(66+66+78+429+2145)q^4+(4\times1+4\times143+2\times780+3432)q^5+\cdots$$

and

$$12q+220q^2+(780+12+12)q^3+(780+780+12+12+12)q^4+(4\times220+2\times572+4928)q^5+\cdots.$$

(Some of the decompositions are ambiguous; for example $1+143$ could be $66+78$. We have chosen $1+143$ instead of $66+78$ for a reason that will appear in a moment.) At first sight this seems to rule out a vertex algebra structure on any lift to characteristic 0, as the homogeneous components V_n, $V_{n+1},\ldots$ have to increase. However we can get around this by rescaling the grading by a factor of $1/3$. The data above seem consistent with the existence of a $\frac{1}{3}\mathbf{Z}$-graded vertex superalgebra $W=\sum_{3n\in\mathbf{Z}}W_n$ such that the mod

3 reduction of W_n is V_{3n}. The vertex superalgebra W is probably acted on by $6.Suz$, with the element of order 3 in the center acting as ω^{3n} on W_n.

There is a second modular vertex super algebra acted on by each of the groups $2.Suz$, $2.HJ$, $2.A_7$, and $2.A_4$, which can be constructed as $\hat{H}^*(g, V_\Lambda)$ where g is a fixed point free automorphism of order 3, 5, 7, or 13 of the Leech lattice Λ. The ordinary part of this superalgebra is the same as the ordinary part of the superalgebra constructed from the monster vertex algebra, but its super part is different; for example, its degree 0 piece does not vanish.

7. Open problems and conjectures.

For more open problems in this area see section 7 of [B-R].

At the end of section 6 we saw that the vertex superalgebra of an element of type $3B$ looks as if it might be the reduction mod 3 of a superalgebra in characteristic 0. More generally something similar seems to happen for some other elements of the monster, as described in the following conjecture.

Conjecture. *For each element $g \in M$ of order n there is a $\frac{1}{n}\mathbf{Z}$-graded super module ${}^g\hat{V} = \sum_{i\in\frac{1}{n}\mathbf{Z}} V_i$ over the ring $\mathbf{Z}[e^{2\pi i/n}]$ with the following properties.*

1. *${}^g\hat{V}$ is often a vertex superalgebra with a conformal vector and a self dual invariant symmetric bilinear form. (For example, it seems to be a vertex algebra when g is of type $1A$, $2B$, $3B$, $5B$, $7B$, $13B$, or $3C$, but possibly not if g has type $2A$ or $3A$.)*
2. *${}^1\hat{V}$ is a self dual integral form of the monster vertex algebra (probably the one conjecturally described in lemma 7.1 of [B-R]).*
3. *If g has prime order then the reduction mod p of ${}^g\hat{V}$ is the modular vertex superalgebra ${}^gV$ associated to g (at least when its grading is multiplied by n.)*
4. *If all coefficients of the Hauptmodul of g are nonnegative then the super part of ${}^g\hat{V}$ vanishes. In this case ${}^g\hat{V}$ should be the same as the "twisted sector" of g that appears in Norton's generalized moonshine conjectures [N]. (But notice that if the Hauptmodul of g has negative coefficients then ${}^g\hat{V}$ is certainly not the twisted sector of g; for example, ${}^g\hat{V}$ is a super vector space and has a nontrivial piece of degree 0.)*
5. *${}^g\hat{V}$ is acted on by a central extension $(\mathbf{Z}/n\mathbf{Z}).C_M(g)$ of $C_M(g)$ by $\mathbf{Z}/n\mathbf{Z}$. This central extension contains an element of order 1 or 2 acting as -1 on the super part of ${}^g\hat{V}$ and as 1 on the ordinary part.*
6. *If g and h are commuting elements of M of coprime orders and $\hat{g}$ is the lift of g to the central extension of $C_M(h)$ with the same order as g, then*
$$\mathrm{Tr}(\hat{g}|{}^h\hat{V}) = \mathrm{Tr}(gh|{}^1\hat{V}).$$
(This identity is often false if the orders of g and h are not coprime.)
7. *If g and h commute then $\hat{H}^*(\hat{g}, {}^h\hat{V})$ is (essentially) the reduction mod $|g|$ of ${}^{gh}V$. (Actually this is not quite correct as stated, because the reduction mod $|g|$ will be defined over some nilpotent extension of $\mathbf{Z}/|g|\mathbf{Z}$, and we should then take the tensor product over this ring with $\mathbf{Z}/|g|\mathbf{Z}$.)*

These conjectures are similar to Norton's generalized moonshine conjectures [N] because both associate some graded space with each element g of the monster. However they treat the problem of negative coefficients in the Hauptmodul $T_g(\tau)$ in a different way: Norton's conjectures use the fact that $T_g(-1/\tau)$ has nonnegative coefficients, which are supposed to be the dimensions of some vector spaces, while the conjectures above allow super vector spaces which may have negative super dimension. Perhaps both Norton's conjectures and the conjectures above are special cases of some conjecture which associates a $\mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}$-graded super space acted on by a central extension $\mathbf{Z}/n\mathbf{Z}.C_M(g)$ to each element $g \in M$ of order n.

We finish by listing a few other questions.

1. Is the group $C_M(g)/O_p(C_M(g))$ the full automorphism group of the modular vertex algebra ${}^gV$? If g has type $3C$ it seems possible that it is not, and the full automorphism group may be $E_8(\mathbf{F}_3)$. Notice that if g has type $2B$ then it is not yet known that $O_2(C_M(g))$ acts trivially on ${}^gV$.
2. Replace the revolting argument of section 4 that involves computer calculations with a more conceptual proof. More generally, suppose we have coefficients $c(n)$ of some function such that the coefficient of $r^m q^n$ in
$$r^{-1}\prod_{m>0, n\in\mathbf{Z}}(1-r^m q^n)^{c(mn)}$$

vanishes whenever $(mn, N) = 1$ for some fixed integer N. What does this imply about the coefficients $c(n)$? (There are many solutions to this given by taking the $c(n)$'s to be the coefficients of some Hauptmodul of level N.)

3. What happens when we apply the argument of section 3 to some automorphism of prime order of the fake monster Lie algebra? Do the infinite products we get represent interesting functions?
4. In [B section 10] it was suggested that there should be Lie superalgebras associated with elements g of the monster, but it was not clear what the dimensions of homogeneous components of these superalgebras should be in the cases when the coefficients of the Hauptmodul of g were neither all positive nor alternating in sign. The results of section 6 suggest what to do in this case: the sum of the dimensions of the ordinary and super parts should also be given by the coefficients of another Hauptmodul.

References.

[BBCO] C. Batut, D. Bernardi, H. Cohen, M. Olivier, "User's guide to PARI-GP". This guide and the PARI programs can be obtained by anonymous ftp from hensel.mathp7.jussieu.fr in the directory /dist/pari.

[B] R. E. Borcherds, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math. 109, 405-444 (1992).

[B-R] R. E. Borcherds, A. J. E. Ryba, Modular Moonshine II, Duke Math Journal Vol. 83 No. 2, 435-459, 1996.

[CG] C. J. Cummins, T. Gannon, Modular equations and the genus zero property of moonshine functions, Invent. Math. 129, 413-443 1997.

[DM] C. Dong and G. Mason, On the construction of the moonshine module as a $\mathbf{Z}_p$-orbifold, Mathematical aspects of conformal and topological field theories and quantum groups (South Hadley, MA, 1992), 37–52, Contemp. Math., 175, Amer. Math. Soc., Providence, RI, 1994.

[FLM] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[Mac] I. G. Macdonald, "Symmetric functions and Hall polynomials", Oxford University press, 1979.

[Ma] Y. Martin, On modular invariance of completely replicable functions. Moonshine, the Monster, and related topics (South Hadley, MA, 1994), 263–286, Contemp. Math., 193, Amer. Math. Soc., Providence, RI, 1996.

[M] P. Montague, Third and Higher Order NFPA Twisted Constructions of Conformal Field Theories from Lattices, Nuclear Phys. B 441 (1995), no. 1-2, 337–382.

[N] S. P. Norton, Generalized Moonshine, Proc. Symp. Pure Math. 47 (1987) p. 208-209.

[R] A. J. E. Ryba, Modular Moonshine?, In "Moonshine, the Monster, and related topics", edited by Chongying Dong and Geoffrey Mason. Contemporary Mathematics, 193. American Mathematical Society, Providence, RI, 1996. 307-336.

WIKIPÉDIA
L'encyclopédie libre

Monstrous moonshine

En mathématiques, **monstrous moonshine** est un terme anglais conçu par John Horton Conway et Simon P. Norton en 1979, utilisé pour décrire la connexion, alors totalement inattendue, entre le groupe Monstre *M* et les formes modulaires (en particulier la fonction *j*).

Précisément, Conway et Norton, suivant une observation initiale de John McKay, trouvèrent que le développement de Fourier de $j(\tau)$ (suite A000521 de l'OEIS, où τ désigne le quotient des demi-périodes `(en)`) pouvait être exprimé en termes de combinaisons linéaires des dimensions des représentations irréductibles de *M* (suite A001379 de l'OEIS)

$$j(\tau) = \frac{1}{q} + 744 + 196884\,q + 21493760\,q^2 + 864299970\,q^3 + \cdots$$

où $q = e^{2\pi i\tau}$ et

$$\begin{aligned} 1 &= 1 \\ 196884 &= 196883 + 1 \\ 21493760 &= 21296876 + 196883 + 1 \\ 864299970 &= 842609326 + 21296876 + 2 \cdot 196883 + 2 \cdot 1 \\ &\vdots \end{aligned}$$

Conway et Norton formulèrent des conjectures concernant les fonctions $j_g(q)$ obtenues en remplaçant les traces sur l'élément neutre par les traces sur d'autres éléments *g* de *M*. La partie la plus saisissante de ces conjectures est que toutes ces fonctions sont de genre zéro. En d'autres termes, si G_g est le sous-groupe de $SL_2(\mathbb{R})$ qui fixe $j_g(q)$, alors le quotient du demi-plan supérieur du plan complexe par G_g est une sphère privée d'un nombre fini de points, correspondant aux formes paraboliques de G_g.

Il s'avère que derrière *monstrous moonshine* se trouve une certaine théorie des cordes ayant le groupe Monstre comme groupe de symétries ; les conjectures faites par Conway et Norton furent démontrées par Richard Ewen Borcherds en 1992 en utilisant le théorème de Goddard-Thorn `(en)` issu de la théorie des cordes, ainsi que la théorie des algèbres vertex et des algèbres de Kac-Moody généralisées `(en)`. Borcherds reçut la médaille Fields pour son travail, et des connexions supplémentaires entre *M* et la fonction *j* furent découvertes ultérieurement.

Sommaire

Versions formelles des conjectures de Conway et Norton
Le module Monstre
La démonstration de Borcherds
Pourquoi « monstrous moonshine » ?
Références
Lien externe

Versions formelles des conjectures de Conway et Norton

La première conjecture faite par Conway et Norton fut ce que l'on appela la « conjecture *moonshine* » ; elle établit qu'il existe un *M*-module gradué de dimension infinie

$$V = \bigoplus_{m \geq -1} V_m$$

avec $\dim(V_m) = c(m)$ pour tout *m*, où

$$j(q) = \sum_{m \geq -1} c_m q^m .$$

De ceci, il s'ensuit que chaque élément *g* de *M* agit sur chaque V_m et possède une valeur de caractère

$$\chi_m(g) = \operatorname{tr}(g|_{V_m})$$

qui peut être utilisée pour construire la **série de McKay-Thompson** de *g* :

$$T_g(q) = \sum_{m \geq -1} \chi_m(g) q^m .$$

La deuxième conjecture de Conway et Norton établit ensuite qu'avec *V* comme ci-dessus, pour chaque élément *g* de *M*, il existe un sous-groupe *K* de $PSL_2(\mathbb{R})$, de genre zéro, commensurable avec le groupe modulaire Γ = $PSL_2(\mathbf{Z})$, et tel que $T_g(q)$ soit la fonction modulaire principale normalisée pour *K*.

Le module Monstre

Il fut montré plus tard par A. O. L. Atkin, Paul Fong et Frederic L. Smith en utilisant des calculs informatiques qu'il existe en effet une représentation graduée de dimension infinie du groupe Monstre dont les séries de McKay-Thompson sont précisément les Hauptmoduls trouvés par Conway et Norton, Igor Frenkel **(en)**, James Lepowsky **(en)** et Arne Meurman **(en)** construisirent explicitement cette représentation en utilisant les opérateurs vertex. Le module résultant est appelé le module Monstre, ou algèbre vertex Monstre **(en)**.

La démonstration de Borcherds

La démonstration de Richard Ewen Borcherds de la conjecture de Conway et Norton peut être séparée en cinq étapes majeures comme ce qui suit :

1. On construit une algèbre vertex *V* qui est une algèbre graduée fournissant les représentations moonshine sur *M*, et on vérifie que le module Monstre possède une structure d'algèbre vertex invariante sous l'action de *M*.
2. On construit une algèbre de Lie $\mathcal{M}$ à partir de *V* en utilisant le théorème de Goddard-Thorn ; ceci est une algèbre de Lie Kac-Moody généralisée.
3. On calcule pour $\mathcal{M}$ une équations des dénominateurs **(en)**, reliée aux coefficients de $j(q)$.
4. On calcule des équations tordues des dénominateurs, qui sont reliées de manière similaire aux séries $T_g(q)$.
5. Toutes ces équations sont utilisées pour déterminer les nombres c en utilisant les opérateurs

5. Toutes ces équations sont utilisées pour déterminer les nombres c_m, en utilisant les opérateurs de Hecke **(en)**, l'homologie d'algèbre de Lie **(en)** et les opérations d'Adams **(en)**.

Ainsi, la démonstration est achevée. Borcherds fut plus tard cité comme ayant dit « j'étais sur la Lune lorsque j'ai démontré la conjecture *moonshine* (clair de lune) » et « Je me demande quelquefois si c'est ce que l'on ressent lorsque l'on prend certaines drogues. Je ne le sais pas actuellement, comme je n'ai pas testé cette théorie personnelle. »

Pourquoi « monstrous moonshine » ?

Le terme « *monstrous moonshine* » fut inventé par Conway, qui, lorsque John McKay lui dit à la fin des années 1970 que les coefficients de q (concrètement 196 884) étaient précisément la dimension de l'algèbre de Griess **(en)** (et ainsi exactement un de plus que le degré de la plus petite représentation fidèle complexe du groupe Monstre), répondit que ceci était « *moonshine* » (au sens de « dingue » ou « idées folles »). Ainsi, le terme fait référence non seulement au groupe Monstre M, mais aussi à la folie perçue concernant la relation compliquée entre M et la théorie des fonctions modulaires.

Cependant, « *moonshine* » est aussi un mot d'argot pour un whisky distillé illégalement, et en fait, le nom peut être expliqué à la lumière de ceci. Le groupe Monstre fut étudié dans les années 1970 par les mathématiciens Jean-Pierre Serre, Andrew Ogg **(en)** et John Griggs Thompson ; ils étudièrent le quotient du plan hyperbolique **(en)** par les sous-groupes de $SL(2,\mathbf{R})$, en particulier le normalisateur $\Gamma_0(p)^+$ du sous-groupe de congruences **(en)** $\Gamma_0(p)$ dans $SL(2,\mathbf{R})$. Ils trouvèrent que la surface de Riemann obtenue en prenant le quotient du plan hyperbolique par $\Gamma_0(p)^+$ est de genre zéro si et seulement si p est 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 41, 47, 59 ou 71 (c'est-à-dire, un nombre premier super-singulier), et lorsque Ogg entendit plus tard parler du groupe Monstre et remarqua que ces nombres étaient précisément les facteurs premiers de la taille de M, il prépara un article offrant une bouteille de whisky Jack Daniel's à quiconque pouvait expliquer ce fait[1].

Références

1. Andrew P. Ogg, *Automorphismes de courbes modulaires.* Seminaire Delange-Pisot-Poitou. Theorie des nombres, tome 16, no. 1 (1974–1975), exp. no. 7, p. 7

- **(en)** Cet article est partiellement ou en totalité issu de l'article de Wikipédia en anglais intitulé « Monstrous moonshine (https://en.wikipedia.org/wiki/Monstrous_moonshine?oldid=80849162) » (voir la liste des auteurs (https://en.wikipedia.org/wiki/Monstrous_moonshine?action=history)).
- Jacques Tits, Le module du « Moonshine » (http://www.numdam.org/numdam-bin/fitem?id=SB_1986-1987__29__285_0), Séminaire Bourbaki, 29 (1986-1987), Exposé No. 684, p. 285-303.
- **(en)** John Horton Conway et Simon P. Norton, *Monstrous Moonshine*, Bull. London Math. Soc. 11, 308–339, 1979.
- **(en)** I. B. Frenkel, J. Lepowsky et A. Meurman, *Vertex Operator Algebras and the Monster*, Pure and Applied Math., Vol. 134, Academic Press, 1988.
- **(en)** Richard Ewen Borcherds, *Monstrous Moonshine and Monstrous Lie Superalgebras*, Invent. Math. 109, 405–444, 1992 [lire en ligne (http://math.berkeley.edu/~reb/papers/)].
- **(en)** Terry Gannon, *Monstrous Moonshine: The first twenty-five years*, 2004 [lire en ligne (https://arxiv.org/abs/math.QA/0402345)].
- **(en)** Terry Gannon, *Monstrous Moonshine and the Classification of Conformal Field Theories*, réimpr. dans *Conformal Field Theory, New Non-Perturbative Methods in String and Field Theory*, (2000) Yavuz Nutku, Cihan Saclioglu, Teoman Turgut, eds. Perseus Publishing, Cambridge Mass (ISBN 0-7382-0204-5) (fournit des exposés introductifs aux applications en physique).

Lien externe

(en) *Moonshine Bibliography (https://web.archive.org/web/20061205032152/http://cicma.mathstat.concordia.ca/faculty/cummins/moonshine.refs.html)*

Ce document provient de « https://fr.wikipedia.org/w/index.php?title=Monstrous_moonshine&oldid=180738689 ».

La dernière modification de cette page a été faite le 10 mars 2021 à 16:03.

Droit d'auteur : les textes sont disponibles sous licence Creative Commons attribution, partage dans les mêmes conditions ; d’autres conditions peuvent s’appliquer. Voyez les conditions d’utilisation pour plus de détails, ainsi que les crédits graphiques. En cas de réutilisation des textes de cette page, voyez comment citer les auteurs et mentionner la licence.
Wikipedia® est une marque déposée de la Wikimedia Foundation, Inc., organisation de bienfaisance régie par le paragraphe 501(c)(3) du code fiscal des États-Unis.

arXiv:1503.01472v2 [math.RT] 26 Nov 2015

Proof of the Umbral Moonshine Conjecture

John F. R. Duncan, Michael J. Griffin and Ken Ono

2015 November 25

Abstract

The Umbral Moonshine Conjectures assert that there are infinite-dimensional graded modules, for prescribed finite groups, whose McKay-Thompson series are certain distinguished mock modular forms. Gannon has proved this for the special case involving the largest sporadic simple Mathieu group. Here we establish the existence of the umbral moonshine modules in the remaining 22 cases.

---

[*] *MSC2010:* 11F22, 11F37.

[*] The authors thank the NSF for its support. The first author also thanks the Simons Foundation (#316779), and the third author thanks the A. G. Candler Fund. The authors thank Miranda Cheng, Jeff Harvey, Michael Somos and the anonymous referee for helpful comments and corrections.

1 Introduction and Statement of Results

Monstrous moonshine relates distinguished modular functions to the representation theory of the Monster, $\mathbb{M}$, the largest sporadic simple group. This theory was inspired by the famous observations of McKay and Thompson in the late 1970s [18,51] that

$$\begin{aligned} 196884 &= 1 + 196883, \\ 21493760 &= 1 + 196883 + 21296876. \end{aligned}$$

The left hand sides here are familiar as coefficients of Klein's modular function (note $q := e^{2\pi i\tau}$),

$$J(\tau) = \sum_{n=-1}^{\infty} c(n)q^n := j(\tau) - 744 = q^{-1} + 196884q + 21493760q^2 + \dots.$$

The sums on the right hand sides involve the first three numbers arising as dimensions of irreducible representations of $\mathbb{M}$,

$$1,\ 196883,\ 21296876,\ 842609326,\ \dots,\ 258823477531055064045234375.$$

Thompson conjectured that there is a graded infinite-dimensional $\mathbb{M}$-module

$$V^\natural = \bigoplus_{n=-1}^{\infty} V_n^\natural,$$

satisfying $\dim(V_n^\natural) = c(n)$. For $g \in \mathbb{M}$, he also suggested [50] to consider the graded-trace functions

$$T_g(\tau) := \sum_{n=-1}^{\infty} \mathrm{tr}(g|V_n^\natural)q^n,$$

now known as the *McKay-Thompson series*, that arise from the conjectured $\mathbb{M}$-module $V^\natural$. Using the character table for $\mathbb{M}$, it was observed [18,50] that the first few coefficients of each $T_g(\tau)$ coincide with those of a generator for the function field of a discrete group $\Gamma_g < SL_2(\mathbb{R})$, leading Conway and Norton [18] to their famous *Monstrous Moonshine Conjecture*: This is the claim that for each $g \in \mathbb{M}$ there is a specific *genus zero* group Γ_g such that $T_g(\tau)$ is the unique normalized *hauptmodul* for Γ_g, i.e., the unique Γ_g-invariant holomorphic function on $\mathbb{H}$ which satisfies $T_g(\tau) = q^{-1} + O(q)$ as $\Im(\tau) \to \infty$, and remains bounded near any cusp not equivalent to the infinite one.

In a series of ground-breaking works, Borcherds introduced vertex algebras [2], and generalized Kac–Moody Lie algebras [3, 4], and used these notions to prove [5] the Monstrous Moonshine Conjecture of Conway and Norton. He confirmed the conjecture for the module $V^\natural$ constructed by Frenkel, Lepowsky, and Meurman [30–32] in the early 1980s. These results provide much more than the predictions of monstrous moonshine. The $\mathbb{M}$-module $V^\natural$ is a vertex operator algebra, one whose automorphism group is precisely $\mathbb{M}$. The construction of Frenkel, Lepowsky and Meurman can be regarded as one of the first examples of an *orbifold conformal field theory*. (Cf. [23].) Here the orbifold in question is the quotient $\left(\mathbb{R}^{24}/\Lambda_{24}\right)/(\mathbb{Z}/2\mathbb{Z})$, of the 24-dimensional torus $\Lambda_{24} \otimes_{\mathbb{Z}} \mathbb{R}/\Lambda_{24} \simeq \mathbb{R}^{24}/\Lambda_{24}$ by the Kummer involution $x \mapsto -x$, where Λ_{24} denotes the Leech lattice.

We refer to [24,32,35,36] for more on monstrous moonshine.

In 2010, Eguchi, Ooguri, and Tachikawa reignited moonshine with their observation [28] that dimensions of some representations of M_{24}, the largest sporadic simple Mathieu group (cf. e.g. [20, 21]), are multiplicities of superconformal algebra characters in the K3 elliptic genus. This observation suggested a manifestation of moonshine for M_{24}: Namely, there should be an infinite-dimensional graded M_{24}-module whose McKay-Thompson series are holomorphic parts of *harmonic Maass forms*, the so-called *mock modular forms*. (See [45, 54, 55] for introductory accounts of the theory of mock modular forms.)

Following the work of Cheng [10], Eguchi and Hikami [27], and Gaberdiel, Hohenegger, and Volpato [33, 34], Gannon established the existence of this infinite-dimensional graded M_{24}-module in [37].

It is natural to seek a general mathematical and physical setting for these results. Here we consider the mathematical setting, which develops from the close relationship between the monster group $\mathbb{M}$ and the Leech lattice Λ_{24}. Recall (cf. e.g. [20]) that the Leech lattice is even, unimodular, and positive-definite of rank 24. It turns out that M_{24} is closely related to another such lattice. Such observations led Cheng, Duncan and Harvey to further instances of moonshine within the setting of even unimodular positive-definite lattices of rank 24. In this way they arrived at the *Umbral Moonshine Conjectures* (cf. §5 of [15], §6 of [16], and §2 of [17]), predicting the existence of 22 further, graded infinite-dimensional modules, relating certain finite groups to distinguished mock modular forms.

To explain this prediction in more detail we recall Niemeier's result [43] that there are 24 (up to isomorphism) even unimodular positive-definite lattices of rank 24. The Leech lattice is the unique one with no root vectors (i.e. lattice vectors with norm-square 2), while the other 23 have root systems with full rank, 24. These *Niemeier root systems* are unions of simple simply-laced root systems with the same Coxeter numbers, and are given explicitly as

$$\begin{gathered} A_1^{24},\ A_2^{12},\ A_3^{8},\ A_4^{6},\ A_6^{4},\ A_{12}^{2}, \\ A_5^4D_4,\ A_7^2D_5^2,\ A_8^3,\ A_9^2D_6,\ A_{11}D_7E_6,\ A_{15}D_9,\ A_{17}E_7,\ A_{24}, \\ D_4^6,\ D_6^4,\ D_8^3,\ D_{10}E_7^2,\ D_{12}^2,\ D_{16}E_8,\ D_{24}, E_6^4,\ E_8^3, \end{gathered} \tag{1.1}$$

in terms of the standard ADE notation. (Cf. e.g. [20] or [39] for more on root systems.)

For each Niemeier root system X let N^X denote the corresponding unimodular lattice, let W^X denote the (normal) subgroup of $\mathrm{Aut}(N^X)$ generated by reflections in roots, and define the *umbral group* of X by setting

$$G^X := \mathrm{Aut}(N^X)/W^X. \tag{1.2}$$

(See §A.1 for explicit descriptions of the groups G^X.)

Let m^X denote the Coxeter number of any simple component of X. An association of distinguished $2m^X$-vector-valued mock modular forms $H_g^X(\tau) = (H_{g,r}^X(\tau))$ to elements $g \in G^X$ is described and analyzed in [15–17].

For $X = A_1^{24}$ we have $G^X \simeq M_{24}$ and $m^X = 2$, and the functions $H_{g,1}^X(\tau)$ are precisely the mock modular forms assigned to elements $g \in M_{24}$ in the works [10, 27, 33, 34] mentioned above. Generalizing the M_{24} moonshine initiated by Eguchi, Ooguri and Tachikawa, we have the following conjecture of Cheng, Duncan and Harvey (cf. §2 of [17] or §9.3 of [24]).

Conjecture (Umbral Moonshine Modules)**.** Let X be a Niemeier root system X and set $m := m^X$. There is a naturally defined bi-graded infinite-dimensional G^X-module

$$\check{K}^X = \bigoplus_{r\in I^X} \bigoplus_{\substack{D\in\mathbb{Z},\ D\leq 0,\\ D=r^2 \pmod{4m}}} \check{K}^X_{r,-D/4m} \tag{1.3}$$

such that the vector-valued mock modular form $H^X_g = (H^X_{g,r})$ is a McKay-Thompson series for $\check{K}^X$ related[1] to the graded trace of g on $\check{K}^X$ by

$$H^X_{g,r}(\tau) = -2q^{-1/4m}\delta_{r,1} + \sum_{\substack{D\in\mathbb{Z},\ D\leq 0,\\ D=r^2 \pmod{4m}}} \operatorname{tr}(g|\check{K}^X_{r,-D/4m})q^{-D/4m} \tag{1.4}$$

for $r \in I^X$.

In (1.3) and (1.4) the set $I^X \subset \mathbb{Z}/2m\mathbb{Z}$ is defined in the following way. If X has an A-type component then $I^X := \{1, 2, 3, \ldots, m-1\}$. If X has no A-type component but does have a D-type component then $m = 2 \mod 4$, and $I^X := \{1, 3, 5, \ldots, m/2\}$. The remaining cases are $X = E_6^4$ and $X = E_8^3$. In the former of these, $I^X := \{1, 4, 5\}$, and in the latter case $I^X := \{1, 7\}$.

Remark. The functions $H^X_g(\tau)$ are defined explicitly in §B.3. An alternative description in terms of Rademacher sums is given in §B.4.

Here we prove the following theorem.

Theorem 1.1. *The umbral moonshine modules exist.*

Two remarks.

1) Theorem 1.1 for $X = A_1^{24}$ is the main result of Gannon's work [37].

2) The vector-valued mock modular forms $H^X = (H^X_{g,r})$ have "minimal" *principal parts*. This minimality is analogous to the fact that the original McKay-Thompson series $T_g(\tau)$ for the Monster are hauptmoduln, and plays an important role in our proof.

Example. Many of Ramanujan's mock theta functions [46] are components of the vector-valued umbral McKay-Thompson series $H^X_g = (H^X_{g,r})$. For example, consider the root system $X = A_2^{12}$, whose umbral group is a double cover $2.M_{12}$ of the sporadic simple Mathieu group M_{12}. In terms of Ramanujan's 3rd order mock theta functions

---

[1] In the statement of Conjecture 6.1 of [16] the function $H^X_{g,r}$ in (1.4) is replaced with $3H^X_{g,r}$ in the case that $X = A_8^3$. This is now known to be an error, arising from a misspecification of some of the functions H^X_g for $X = A_8^3$. Our treatment of the case $X = A_8^3$ in this work reflects the corrected specification of the corresponding H^X_g which is described and discussed in detail in [17].

$$\begin{aligned}
f(q) &= 1+\sum_{n=1}^{\infty}\frac{q^{n^2}}{(1+q)^2(1+q^2)^2\cdots(1+q^n)^2},\\
\phi(q) &= 1+\sum_{n=1}^{\infty}\frac{q^{n^2}}{(1+q^2)(1+q^4)\cdots(1+q^{2n})},\\
\chi(q) &= 1+\sum_{n=1}^{\infty}\frac{q^{n^2}}{(1-q+q^2)(1-q^2+q^4)\cdots(1-q^n+q^{2n})}\\
\omega(q) &= \sum_{n=0}^{\infty}\frac{q^{2n(n+1)}}{(1-q)^2(1-q^3)^2\cdots(1-q^{2n+1})^2},\\
\rho(q) &= \sum_{n=0}^{\infty}\frac{q^{2n(n+1)}}{(1+q+q^2)(1+q^3+q^6)\cdots(1+q^{2n+1}+q^{4n+2})},
\end{aligned}$$

we have that

$$\begin{aligned}
H^X_{2B,1}(\tau)=H^X_{2C,1}(\tau)=H^X_{4C,1}(\tau) &= -2q^{-\frac{1}{12}}\cdot f(q^2),\\
H^X_{6C,1}(\tau)=H^X_{6D,1}(\tau) &= -2q^{-\frac{1}{12}}\cdot \chi(q^2),\\
H^X_{8C,1}(\tau)=H^X_{8D,1}(\tau) &= -2q^{-\frac{1}{12}}\cdot \phi(-q^2),\\
H^X_{2B,2}(\tau)=-H^X_{2C,2}(\tau) &= -4q^{\frac{2}{3}}\cdot \omega(-q),\\
H^X_{6C,2}(\tau)=-H^X_{6D,2}(\tau) &= 2q^{\frac{2}{3}}\cdot \rho(-q).
\end{aligned}$$

See §5.4 of [16] for more coincidences between umbral McKay-Thompson series and mock theta functions identified by Ramanujan almost a hundred years ago.

Our proof of Theorem 1.1 involves the explicit determination of each G^X-module $\check{K}^X$ by computing the multiplicity of each irreducible component for each homogeneous subspace. It guarantees the existence and uniqueness of a $\check{K}^X$ which is compatible with the representation theory of G^X and the Fourier expansions of the vector-valued mock modular forms $H^X_g(\tau)=(H^X_{g,r}(\tau))$.

At first glance our methods do not appear to shed light on any deeper algebraic properties of the $\check{K}^X$, such as might correspond to the vertex operator algebra structure on $V^\natural$, or the monster Lie algebra introduced by Borcherds in [5]. However, we do determine, and utilize, specific recursion relations for the coefficients of the umbral McKay-Thompson series which are analogous to the replicability properties of monstrous moonshine formulated by Conway and Norton in §8 of [18] (cf. also [1]). More specifically, we use recent work [41] of Imamoḡlu, Raum and Richter, as generalized [42] by Mertens, to obtain such recursions. These results are based on the process of *holomorphic projection*.

Theorem 1.2. *For each $g\in G^X$ and $0<r<m$, the mock modular form $H^X_{g,r}(\tau)$ is replicable in the mock modular sense.*

A key step in Borcherds' proof [5] of the monstrous moonshine conjecture is the reformulation of replicability in Lie theoretic terms. We may speculate that the *mock modular replicability* utilized in this work will ultimately admit an analogous algebraic interpretation. Such a result remains an important goal for future work.

In the statement of Theorem 1.2, replicable means that there are explicit recursion relations for the coefficients of the vector-valued mock modular form in question. For example, we recall the recurrence formula for Ramanujan's third order mock theta function $f(q) = \sum_{n=0}^{\infty} c_f(n)q^n$ that was obtained recently by Imamoḡlu, Raum and Richter [41]. If $n \in \mathbb{Q}$, then let

$$\sigma_1(n) := \begin{cases} \sum_{d|n} d & \text{if } n \in \mathbb{Z}, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathrm{sgn}^+(n) := \begin{cases} \mathrm{sgn}(n) & \text{if } n \neq 0, \\ 1 & \text{if } n = 0, \end{cases}$$

and then define

$$d(N, \widetilde{N}, t, \widetilde{t}) := \mathrm{sgn}^+(N) \cdot \mathrm{sgn}^+(\widetilde{N}) \cdot \left(|N+t| - |\widetilde{N} + \widetilde{t}| \right).$$

Then for positive integers n, we have that

$$\sum_{\substack{m \in \mathbb{Z} \\ 3m^2+m \leq 2n}} \left(m + \frac{1}{6}\right) c_f\left(n - \frac{3}{2}m^2 - \frac{1}{2}m\right) = \frac{4}{3}\sigma(n) - \frac{16}{3}\sigma\left(\frac{n}{2}\right) - 2\sum_{\substack{a,b \in \mathbb{Z} \\ 2n=ab}} d\left(N, \widetilde{N}, \frac{1}{6}, \frac{1}{6}\right),$$

where $N := \frac{1}{6}(-3a+b-1)$ and $\widetilde{N} := \frac{1}{6}(3a+b-1)$, and the sum is over integers a, b for which $N, \widetilde{N} \in \mathbb{Z}$. This is easily seen to be a recurrence relation for the coefficients $c_f(n)$. The replicability formulas for all of the $H^X_{g,r}(\tau)$ are similar (although some of these relations are slightly more complicated and involve the coefficients of weight 2 cusp forms).

It is important to emphasize that, despite the progress which is represented by our main results, Theorems 1.1 and 1.2, the following important question remains open in general.

Question. *Is there a "natural" construction of $\check{K}^X$? Is $\check{K}^X$ equipped with a deeper algebra structure as in the case of the monster module $V^\natural$ of Frenkel, Lepowsky and Meurman?*

We remark that this question has been answered positively, recently, in one special case: A vertex operator algebra structure underlying the umbral moonshine module $\check{K}^X$ for $X = E_8^3$ has been described explicitly in [25]. See also [14, 26], where the problem of constructing algebraic structures that illuminate the umbral moonshine observations is addressed from a different point of view.

The proof of Theorem 1.1 is not difficult. It is essentially a collection of tedious calculations. We use the theory of mock modular forms and the character table for each G^X (cf. §A.2) to solve for the multiplicities of the irreducible G^X-module constituents of each homogeneous subspace in the alleged G^X-module $\check{K}^X$. To prove Theorem 1.1 it suffices to prove that these multiplicities are non-negative integers. To prove Theorem 1.2 we apply recent work [42] of Mertens on the holomorphic projection of weight $\frac{1}{2}$ mock modular forms, which generalizes earlier work [41] of Imamoḡlu, Raum and Richter.

In §2 we recall the facts about mock modular forms that we require, and we prove Theorem 1.2. We prove Theorem 1.1 in §3. The appendices furnish all the data that our method requires. In particular, the umbral groups G^X are described in detail in §A, and explicit definitions for the mock modular forms $H_g^X(\tau)$ are given in §B.

2 Harmonic Maass forms and Mock modular forms

Here we recall some very basic facts about harmonic Maass forms as developed by Bruinier and Funke [9] (see also [45]).

We begin by briefly recalling the definition of a *harmonic Maass form* of weight $k \in \frac{1}{2}\mathbb{Z}$ and multiplier ν (a generalization of the notion of a Nebentypus). If $\tau = x + iy$ with x and y real, we define the weight k hyperbolic Laplacian by

$$\Delta_k := -y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+iky\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right). \tag{2.1}$$

Suppose Γ is a subgroup of finite index in $SL_2(\mathbb{Z})$ and $k \in \frac{1}{2}\mathbb{Z}$. Then a function $F(\tau)$ which is real-analytic on the upper half of the complex plane is a *harmonic Maass form* of weight k on Γ with multiplier ν if:

(a) The function $F(\tau)$ satisfies the weight k modular transformation,

$$F(\tau)|_k\gamma = \nu(\gamma)F(\tau)$$

for every matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, where $F(\tau)|_k\gamma := F(\gamma\tau)(c\tau+d)^{-k}$, and if $k \in \mathbb{Z}+\frac{1}{2}$, the square root is taken to be the principal branch.

(b) We have that $\Delta_k F(\tau) = 0$,

(c) There is a polynomial $P_F(q^{-1})$ and a constant $c > 0$ such that $F(\tau) - P_F(e^{-2\pi i\tau}) = O(e^{-cy})$ as $\tau \to i\infty$. Analogous conditions are required at each cusp of Γ.

We denote the $\mathbb{C}$-vector space of harmonic Maass forms of a given weight k, group Γ and multiplier ν by $H_k(\Gamma, \nu)$. If no multiplier is specified, we will take

$$\nu_0(\gamma) := \left(\left(\frac{c}{d}\right)\sqrt{\left(\frac{-1}{d}\right)}^{\,-1}\right)^{2k},$$

where $\left(\frac{*}{d}\right)$ is the Kronecker symbol.

2.1 Main properties

The Fourier expansion of a harmonic Maass form F (see Proposition 3.2 of [9]) splits into two components. As before, we let $q := e^{2\pi i\tau}$.

Lemma 2.1. *If $F(\tau)$ is a harmonic Maass form of weight $2-k$ for Γ where $\frac{3}{2} \leq k \in \frac{1}{2}\mathbb{Z}$, then*

$$F(\tau) = F^{+}(\tau) + F^{-}(\tau),$$

where F^{+} is the holomorphic part of F, given by

$$F^{+}(\tau) = \sum_{n \gg -\infty} c_F^{+}(n) q^n$$

where the sum admits only finitely many non-zero terms with $n < 0$, and F^{-} is the nonholomorphic part, given by

$$F^{-}(\tau) = \sum_{n<0} c_F^{-}(n) \Gamma(k-1, 4\pi y|n|) q^n.$$

Here $\Gamma(s, z)$ is the upper incomplete gamma function.

The holomorphic part of a harmonic Maass form is called a *mock modular form.* We denote the space of harmonic Maass forms of weight $2-k$ for Γ and multiplier ν by $H_k(\Gamma, \nu)$. Similarly, we denote the corresponding subspace of holomorphic modular forms by $M_k(\Gamma, \nu)$, and the space of cusp forms by $S_k(\Gamma, \nu)$. The differential operator $\xi_w := 2iy^w \overline{\frac{\partial}{\partial \overline{\tau}}}$ (see [9]) defines a surjective map

$$\xi_{2-k} : H_{2-k}(\Gamma, \nu) \to S_k(\Gamma, \overline{\nu})$$

onto the space of weight k cusp forms for the same group but conjugate multiplier. The *shadow* of a Maass form $f(\tau) \in H_{2-k}(\Gamma, \nu)$ is the cusp form $g(\tau) \in S_k(\Gamma, \overline{\nu})$ (defined, for now, only up to scale) such that $\xi_{2-k} f(\tau) = \frac{g}{||g||}$, where $|| \bullet ||$ denotes the usual Petersson norm.

2.2 Holomorphic projection of weight $\frac{1}{2}$ mock modular forms

As noted above, the modular transformations of a weight $\frac{1}{2}$ harmonic Maass form may be simplified by multiplying by its shadow to obtain a weight 2 nonholomorphic modular form. One can use the theory of holomorphic projections to obtain explicit identities relating these nonholomorphic modular forms to classical quasimodular forms. In this way, we may essentially reduce many questions about the coefficients of weight $\frac{1}{2}$ mock modular forms to questions about weight 2 holomorphic modular forms. The following theorem is a special case of a more general theorem due to Mertens (cf. Theorem 6.3 of [42]). See also [41].

Theorem 2.2 (Mertens)**.** *Suppose $g(\tau)$ and $h(\tau)$ are both theta functions of weight $\frac{3}{2}$ contained in $S_{\frac{3}{2}}(\Gamma, \nu_g)$ and $S_{\frac{3}{2}}(\Gamma, \nu_h)$ respectively, with Fourier expansions*

$$g(\tau) : = \sum_{i=1}^{s} \sum_{n \in \mathbb{Z}} n \chi_i(n) q^{n^2},$$

$$h(\tau) := \sum_{j=1}^{t} \sum_{n \in \mathbb{Z}} n \psi_j(n) q^{n^2},$$

where each χ_i *and* ψ_i *is a Dirichlet character. Moreover, suppose* $h(\tau)$ *is the shadow of a weight* $\frac{1}{2}$ *harmonic Maass form* $f(\tau) \in H_{\frac{1}{2}}(\Gamma, \overline{\nu_h})$*. Define the function*

$$D^{f,g}(\tau) := 2 \sum_{r=1}^{\infty} \sum_{\chi_i, \psi_j} \sum_{\substack{m,n \in \mathbb{Z}^+ \\ m^2 - n^2 = r}} \chi_i(m) \overline{\psi_j(n)} (m-n) q^r.$$

If $f(\tau)g(\tau)$ *has no singularity at any cusp, then* $f^+(\tau)g(\tau) + D^{f,g}(\tau)$ *is a weight* 2 *quasimodular form. In other words, it lies in the space* $\mathbb{C}E_2(\tau) \oplus M_2(\Gamma, \nu_g \overline{\nu_h})$*, where* $E_2(\tau)$ *is the quasimodular Eisenstein series* $E_2(\tau) := 1 - 24 \sum_{n \geq 1} \frac{nq^n}{1-q^n}$.

Two Remarks.
1) These identities give recurrence relations for the weight $\frac{1}{2}$ mock modular form f^+ in terms of the weight 2 quasimodular form which equals $f^+(\tau)g(\tau) + D^{f,g}(\tau)$. The example after Theorem 1.2 for Ramanujan's third order mock theta function f is an explicit example of such a relation.

2) Theorem 2.2 extends to vector-valued mock modular forms in a natural way.

Proof of Theorem 1.2. Fix a Niemeier lattice and its root system X, and let $M = m^X$ denote its Coxeter number. Each $H^X_{g,r}(\tau)$ is the holomorphic part of a weight $\frac{1}{2}$ harmonic Maass form $\widehat{H}^X_{g,r}(\tau)$. To simplify the exposition in the following section, we will emphasize the case that the root system X is of pure A-type. If the root system X is of pure A-type, the shadow function $S^X_{g,r}(\tau)$ is given by $\hat{\chi}^{X_A}_{g,r} S_{M,r}(\tau)$ (see §B.2), where

$$S_{M,r}(\tau) = \sum_{\substack{n \in \mathbb{Z} \\ n \equiv r \pmod{2M}}} n\, q^{\frac{n^2}{4M}},$$

and $\hat{\chi}^{X_A}_{g,r} = \chi^{X_A}_g$ or $\bar{\chi}^{X_A}_g$ depending on the parity of r is the twisted Euler character given in the appropriate table in §A.3, a character of G^X. (If X is not of pure A-type, then the shadow function $S^X_{g,r}(\tau)$ is a linear combination of similar functions as described in §B.2.)

Given X and g, the symbol $n_g|h_g$ given in the corresponding table in §A.3 defines the modularity for the vector-valued function $(\widehat{H}^X_{g,r}(\tau))$. In particular, if the shadow $(S^X_{g,r}(\tau))$ is nonzero, and if for $\gamma \in \Gamma_0(n_g)$ we have that

$$(S^X_{g,r}(\tau))|_{3/2}\gamma = \sigma_{g,\gamma}(S^X_{g,r}(\tau)),$$

then

$$(\widehat{H}^X_{g,r}(\tau))|_{1/2}\gamma = \overline{\sigma_{g,\gamma}}(\widehat{H}^X_{g,r}(\tau)).$$

Here, for $\gamma \in \Gamma_0(n_g)$, we have $\sigma_{g,\gamma} = \nu_g(\gamma)\sigma_{e,\gamma}$ where $\nu_g(\gamma)$ is a multiplier which is trivial on $\Gamma_0(n_g h_g)$. This identity holds even in the case that the shadow $S^X_{g,r}$ vanishes.

The vector-valued function $(H^X_{g,r}(\tau))$ has poles only at the infinite cusp of $\Gamma_0(n_g)$, and only at the component $H^X_{g,r}(\tau)$ where $r = 1$ if X has pure A-type, or at components where $r^2 \equiv 1 \pmod{4M}$ otherwise. These poles may only have order $\frac{1}{4M}$. This implies that the function $(\widehat{H}^X_{g,r}(\tau) S^X_{g,r}(\tau))$ has no pole at any cusp, and is therefore a candidate for an application of Theorem 2.2.

The modular transformation of $S_{M,r}(\tau)$ implies that

$$(\sigma_{e,S})^2 = (\sigma_{e,T})^{4M} = \mathbf{I}$$

where $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, and $\mathbf{I}$ is the identity matrix. Therefore $S^X_{M,r}(\tau)$, viewed as a scalar-valued modular function, is modular on $\Gamma(4M)$, and so $(\widehat{H}^X_{g,r}(\tau) S^X_{g,r}(\tau))$ is a weight 2 nonholomorphic scalar-valued modular form for the group $\Gamma(4M) \cap \Gamma_0(n_g)$ with trivial multiplier.

Applying Theorem 2.2, we obtain a function $F^X_{g,r}(\tau)$—call it the holomorphic projection of $\widehat{H^X_{g,r}}(\tau) S^X_{e,r}(\tau)$—which is a weight 2 quasimodular form on $\Gamma(4M) \cap \Gamma_0(n_g)$. In the case that $S^X_{g,r}(\tau)$ is zero, we substitute $S^X_{e,r}(\tau)$ in its place to obtain a function $\widetilde{F}^X_{g,r}(\tau) = H^X_{g,r}(\tau) S^X_{e,r}(\tau)$ which is a weight 2 holomorphic scalar-valued modular form for the group $\Gamma(4M) \cap \Gamma_0(n_g)$ with multiplier ν_g (alternatively, modular for the group $\Gamma(4M) \cap \Gamma_0(n_g h_g)$ with trivial multiplier).

The function $F^X_{g,r}(\tau)$ may be determined explicitly as the sum of Eisenstein series and cusp forms on $\Gamma(4M) \cap \Gamma_0(n_g h_g)$ using the standard arguments from the theory of holomorphic modular forms (i.e. the “first few” coefficients determine such a form). Therefore, we have the identity

$$F^X_{g,r}(\tau) = H^X_{g,r}(\tau) \cdot S^X_{g,r}(\tau) + D^X_{g,r}(\tau), \tag{2.2}$$

where the function $D^X_{g,r}(\tau)$ is the correction term arising in Theorem 2.2. If X has pure A-type, then

$$D^X_{g,r}(\tau) = (\hat{\chi}^{X_A}_{g,r})^2 \sum_{N=1}^{\infty} \sum_{\substack{m,n \in \mathbb{Z}_+ \\ m^2 - n^2 = N}} \phi_r(m)\phi_r(n)(m-n) q^{\frac{N}{4M}}, \tag{2.3}$$

where

$$\phi_r(\ell) = \begin{cases} \pm 1 & \text{if } \ell \equiv \pm r \pmod{2M} \\ 0 & \text{otherwise.} \end{cases}$$

Suppose $H^X_{g,r}(\tau) = \sum_{n=0}^{\infty} A^X_{g,r}(n) q^{n - \frac{D}{4M}}$ where $0 < D < 4M$ and $D \equiv r^2 \pmod{4M}$, and $F^X_{g,r}(\tau) = \sum_{N=0}^{\infty} B^X_{g,r}(n) q^n$. Then by Theorem 2.2, we find that

$$B^X_{g,r}(N) = \hat{\chi}^{X_A}_{g,r} \sum_{\substack{m \in \mathbb{Z} \\ m \equiv r \pmod{2M}}} m \cdot A^X_{g,r}\left(N + \frac{D - m^2}{4M}\right) + (\hat{\chi}^{X_A}_{g,r})^2 \sum_{\substack{m,n \in \mathbb{Z}^+ \\ m^2 - n^2 = N}} \phi_r(m)\phi_r(n)(m-n). \tag{2.4}$$

The function $F^X_{g,r}(\tau)$ may be found in the following manner. Using the explicit prescriptions for $H^X_{g,r}(\tau)$ given in §B.3 and (2.2) above, we may calculate the first several coefficients of each component. The Eisenstein component is determined by the constant terms at cusps. Since $D^X_{g,r}(\tau)$ (and the corresponding correction terms at other cusps) has no constant term, these are the same as the constant terms of $\widehat{H^X_{g,r}}(\tau) S^X_{g,r}(\tau)$, which are determined by the poles of $\widehat{H^X_{g,r}}$. Call this Eisenstein component $E^X_{g,r}(\tau)$. The cuspidal component can be found by matching the initial coefficients of $F^X_{g,r}(\tau) - E^X_{g,r}(\tau)$.

Once the coefficients $B^X_{g,r}(n)$ are known, equation (2.4) provides a recursion relation which may be used to calculate the coefficients of $H^X_{g,r}(\tau)$. If the shadows $S^X_{g,r}(\tau)$ are zero, then we may apply a similar procedure in order to determine $\widetilde{F}^X_{g,r}(\tau)$. For example, suppose $\widetilde{F}^X_{g,r}(\tau) = \sum_{N=0}^{\infty} \widetilde{B}^X_{g,r}(n)q^n$, and X has pure A-type. Then we find that the coefficients $\widetilde{B}^X_{g,r}(N)$ satisfy

$$\widetilde{B}^X_{g,r}(N) = \hat{\chi}^{X_A}_{g,r} \sum_{\substack{m\in\mathbb{Z} \\ m\equiv r \pmod{2M}}} m \cdot A^X_{g,r}\left(N + \frac{D-m^2}{4M}\right) \tag{2.5}$$

Proceeding in this way we obtain the claimed results. □

3 Proof of Theorem 1.1

Here we prove Theorem 1.1. The idea is as follows. For each Niemeier root system X we begin with the vector-valued mock modular forms $(H^X_g(\tau))$ for $g \in G^X$. We use their q-expansions to solve for the q-series whose coefficients are the alleged multiplicities of the irreducible components of the alleged infinite-dimensional G^X-module

$$\check{K}^X = \bigoplus_{r \pmod{2m}} \bigoplus_{\substack{D\in\mathbb{Z},\, D\le 0, \\ D=r^2 \pmod{4m}}} \check{K}^X_{r,-D/4m}.$$

These q-series turn out to be mock modular forms. The proof requires that we establish that these mock modular forms have non-negative integer coefficients.

Proof of Theorem 1.1. As in the previous section, we fix a root system X and set $M := m^X$, and we emphasize the case when X is of pure A-type.

The umbral moonshine conjecture asserts that

$$H^X_{g,r}(\tau) = \sum_{n=0}^{\infty}\sum_{\chi} m^X_{\chi,r}(n)\chi(g)q^{n-\frac{r^2}{4M}} \tag{3.1}$$

where the second sum is over the irreducible characters of G^X. Here we have rewritten the traces of the graded components $\check{K}^X_{r,n-r^2/4M}$ in 1.4 in terms of the values of the irreducible characters of G^X, where the $m^X_{\chi,r}(n)$ are the corresponding multiplicities. Naturally, if such a $\check{K}^X$ exists, these multiplicities must be non-negative integers for $n > 0$. Similarly, if the mock modular forms $H^X_{g,r}(\tau)$ can be expressed as in 3.1 with $m^X_{\chi,r}(n)$ non-negative integers, then we may construct the umbral moonshine module $\check{K}^X$ explicitly with $\check{K}^X_{r,n-r^2/4M}$ defined as the direct sum of irreducible components with the given multiplicities $m^X_{\chi,r}(n)$.

Let

$$H^X_{\chi,r}(\tau) := \frac{1}{|G^X|}\sum_{g} \overline{\chi(g)}H^X_{g,r}(\tau). \tag{3.2}$$

It turns out that the coefficients of $H^X_{\chi,r}(\tau)$ are precisely the multiplicities $m^X_{\chi,r}(n)$ required so that 3.1 holds: if

$$H^X_{\chi,r}(\tau) = \sum_{n=0}^{\infty} m^X_{\chi,r}(n) q^{n-\frac{r^2}{4M}}, \tag{3.3}$$

then

$$H^X_{g,r}(\tau) = \sum_{n=0}^{\infty} \sum_{\chi} m^X_{\chi,r}(n)\chi(g) q^{n-\frac{r^2}{4M}}.$$

Thus the umbral moonshine conjecture is true if and only if the Fourier coefficients of $H^X_{\chi,r}(\tau)$ are non-negative integers.

To see this fact, we recall the orthogonality of characters. For irreducible characters χ_i and χ_j,

$$\frac{1}{|G^X|} \sum_{g\in G^X} \overline{\chi_i(g)}\chi_j(g) = \begin{cases} 1 & \text{if } \chi_i = \chi_j, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

We also have the relation for g and $h \in G^X$,

$$\sum_{\chi} \overline{\chi_i(g)}\chi_i(h) = \begin{cases} |C_{G^X}(g)| & \text{if } g \text{ and } h \text{ are conjugate,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

Here $|C_{G^X}(g)|$ is the order of the centralizer of g in G^X. Since the order of the centralizer times the order of the conjugacy class of an element is the order of the group, (3.2) and (3.5) together imply the relation

$$H^X_{g,r}(\tau) = \sum_{\chi} \chi(g) H^X_{\chi,r}(\tau),$$

which in turn implies 3.3.

We have reduced the theorem to proving that the coefficients of certain weight $\frac{1}{2}$ mock modular forms are all non-negative integers. For holomorphic modular forms we may answer questions of this type by making use of Sturm's theorem [49] (see also Theorem 2.58 of [44]). This theorem provides a bound B associated to a space of modular forms such that if the first B coefficients of a modular form $f(\tau)$ are integral, then all of the coefficients of $f(\tau)$ are integral. This bound reduces many questions about the Fourier coefficients of modular forms to finite calculations.

Sturm's theorem relies on the finite dimensionality of certain spaces of modular forms, and so it can not be applied directly to spaces of mock modular forms. However, by making use of holomorphic projection we can adapt Sturm's theorem to this setting.

Let $\widehat{H^X_{\chi,r}}(\tau)$ be defined as above. Recall that the transformation matrix for the vector-valued function $\widehat{H^X_{g,r}}(\tau))$ is $\overline{\sigma_{g,\gamma}}$, the conjugate of the transformation matrix for $(S^X_{e,r}(\tau))$ when $\gamma \in \Gamma_0(n_g h_g)$, and $\sigma_{g,\gamma}$ is the identity for $\gamma \in \Gamma(4M)$. Therefore if

$$N^X_{\chi} := \mathrm{lcm}\{n_g h_g \mid g \in G, \chi(g) \neq 0\},$$

then the scalar-valued functions $\widehat{H^X_{\chi,r}}(\tau)$ are modular on $\Gamma(4M) \cap \Gamma_0(N^X_{\chi})$.

Let

$$A_{\chi,r}(\tau) := H^X_{\chi,r}(\tau) S^X_{e,1}(\tau),$$

and let $\tilde{A}_{\chi,r}(\tau)$ be the holomorphic projection of $A_{\chi,r}(\tau)$. Suppose that $H^X_{\chi,r}(\tau)$ has integral coefficients up to some bound B. Formulas for the shadow functions (cf. §B.2) show that the leading coefficient of $S^X_{e,1}(\tau)$ is 1 and has integral coefficients. This implies that the function

$$A_{\chi,r}(\tau) := H^X_{\chi,r}(\tau) S^X_{e,1}(\tau)$$

also has integral coefficients up to the bound B. The shadow of $H^X_{\chi,r}(\tau)$ is given by

$$S^X_{\chi,r}(\tau) := \frac{1}{|G^X|} \sum_g \overline{\chi(g)} S^X_{g,r}(\tau).$$

If X is pure A-type, then $S^X_{g,r}(\tau) = \chi^{X_A}_{g,r} S_{M,r}(\tau) = (\chi'(g) + \chi''(g)) S_{M,r}(\tau)$ for some irreducible characters χ' and χ'', according to §A.3 and §B.2. Therefore,

$$S^X_{\chi,r}(\tau) = \begin{cases} S_{M,r}(\tau) & \text{if } \chi = \chi' \text{ or } \chi'', \\ 0 & \text{otherwise.} \end{cases}$$

When X is not of pure A-type the shadow is some sum of such functions, but in every case has integer coefficients, and so, applying Theorem 2.2 to $A_{\chi,r}(\tau)$, we find that $\tilde{A}_{\chi,r}(\tau)$ also has integer coefficients up to the bound B. In particular, since $\tilde{A}_{\chi,r}(\tau)$ is modular on $\Gamma(4M) \cap \Gamma_0(N^X_\chi)$, then if B is at least the Sturm bound for this group we have that every coefficient of $\tilde{A}_{\chi,r}(\tau)$ is integral. Since the leading coefficient of $S^X_{e,1}(\tau)$ is 1, we may reverse this argument and obtain that every coefficient of $H^X_{\chi,r}(\tau)$ is integral. Therefore, in order to check that $H^X_{\chi,r}(\tau)$ has only integer coefficients, it suffices to check up to the Sturm bound for $\Gamma(4M) \cap \Gamma_0(N_\chi)$. These calculations were carried out using the `sage` mathematical software [47].

The calculations and argument given above shows that the multiplicities $m^X_{\chi,r}(n)$ are all integers. To complete the proof, it suffices to check that they are also are non-negative. The proof of this claim follows easily by modifying step-by-step the argument in Gannon's proof of non-negativity in the M_{24} case [37] (i.e. $X = A_1^{24}$). Here we describe how this is done.

Expressions for the alleged McKay-Thompson series $H^X_{g,r}(\tau)$ in terms of Rademacher sums and unary theta functions are given in §B.4. Exact formulas are known for all the coefficients of Rademacher sums because they are defined by averaging the special function $\mathrm{r}^{[\alpha]}_{1/2}(\gamma, \tau)$ (see (B.114)) over cosets of a specific modular group modulo Γ_∞, the subgroup of translations. Therefore, Rademacher sums are standard Maass-Poincaré series, and as a result we have formulas for each of their coefficients as convergent infinite sums of Kloosterman-type sums weighted by values of the $I_{1/2}$ modified Bessel function. (For example, see [8] or [53] for the general theory, and [12] for the specific case that $X = A_1^{24}$.) More importantly, this means also that the generating function for the multiplicities $m^X_{\chi,r}(n)$ is a weight $\frac{1}{2}$ harmonic Maass form, which in turn means that exact formulas (modulo the unary theta functions) are also available in similar terms. For positive integers n, this then means that (cf. Theorem 1.1 of [8])

$$m^X_{\chi,r}(n) = \sum_\rho \sum_{m<0} \frac{a^X_\rho(m)}{n^{\frac{1}{4}}} \sum_{c=1}^\infty \frac{K^X_\rho(m,n,c)}{c} \cdot \mathbb{I}^X \left(\frac{4\pi\sqrt{|nm|}}{c} \right), \tag{3.6}$$

where the sums are over the cusps ρ of the group $\Gamma_0(N^X_g)$, and finitely many explicit negative rational numbers m. The constants $a^X_\rho(m)$ are essentially the coefficients which describe the generating function in terms of Maass-Poincaré series. Here $\mathbb{I}$ is a suitable normalization and change of variable for the standard $I_{1/2}$ modified Bessel-function.

The Kloosterman-type sums $K_\rho^X(m,n,c)$ are well known to be related to Salié-type sums (for example see Proposition 5 of [40]). These Salié-type sums are of the form

$$S_\rho^X(m,n,c) = \sum_{\substack{x \pmod{c} \\ x^2 \equiv -D(m,n) \pmod{c}}} \epsilon_\rho^X(m,n) \cdot e\left(\frac{\beta^X x}{c}\right),$$

where $\epsilon_\rho^X(m,n)$ is a root of unity, $-D(m,n)$ is a discriminant of a positive definite binary quadratic form, and β^X is a nonzero positive rational number.

These Salié sums may then be estimated using the equidistribution of CM points with discriminant $-D(m,n)$. This process was first introduced by Hooley [38], and it was first applied to the coefficients of weight $\frac{1}{2}$ mock modular forms by Bringmann and Ono [7]. Gannon explains how to make effective the estimates for sums of this shape in §4 of [37], thereby reducing the proof of the M_{24} case of umbral moonshine to a finite calculation. In particular, in equations (4.6-4.10) of [37] Gannon shows how to bound coefficients of the form (3.6) in terms of the Selberg–Kloosterman zeta function, which is bounded in turn in his proof of Theorem 3 of [37]. We follow Gannon's proof *mutatis mutandis*. We find, for most root systems, that the coefficients of each multiplicity generating function are positive beyond the 390th coefficient. In the worst case, for the root system $X = E_8^3$, we find that the coefficients are positive beyond the 1780th coefficient. Moreover, the coefficients exhibit subexponential growth. A finite computer calculation in `sage` has verified the non-negativity of the finitely many remaining coefficients. ☐

Remark. It turns out that the estimates required for proving non-negativity are the worst for the $X = E_8^3$ case which required 1780 coefficients.

A The Umbral Groups

In this section we present the facts about the umbral groups that we have used in establishing the main results of this paper. We recall (from [16]) their construction in terms of Niemeier root systems in §A.1, and we reproduce their character tables (appearing also in [16]) in §A.2. Note that we use the abbreviations $a_n := \sqrt{-n}$ and $b_n := (-1+\sqrt{-n})/2$ in the tables of §A.2.

The root system description of the umbral groups (cf. §A.1) gives rise to certain characters called *twisted Euler characters* which we recall (from [16]) in §A.3. The data appearing in §A.3 plays an important role in §B.2, where we use it to describe the shadows S_g^X of the umbral McKay-Thompson series H_g^X explicitly.

A.1 Construction

As mentioned in §1, there are exactly 24 self-dual even positive-definite lattices of rank 24 up to isomorphism, according to the classification of Niemeier [43] (cf. also [19, 52]). Such a lattice L is determined up to isomorphism by its *root system* $L_2 := \{\alpha \in L \mid \langle \alpha, \alpha \rangle = 2\}$. The unique example without roots is the Leech lattice. We refer to the remaining 23 as the *Niemeier lattices*, and we call a root system X a *Niemeier root system* if it occurs as the root system of a Niemeier lattice.

The simple components of Niemeier root systems are root systems of ADE type, and it turns out that the simple components of a Niemeier root system X all have the same Coxeter number. Define m^X to be the Coxeter number of any simple component of X, and call this the *Coxeter number* of X.

For X a Niemeier root system write N^X for the corresponding Niemeier lattice. The *umbral group* attached to X is defined by setting

$$G^X := \mathrm{Aut}(N^X)/W^X \tag{A.1}$$

where W^X is the normal subgroup of $\mathrm{Aut}(N^X)$ generated by reflections in root vectors.

Observe that G^X acts as permutations on the simple components of X. In general this action is not faithful, so define $\overline{G}^X$ to be the quotient of G^X by its kernel. It turns out that the level of the mock modular form H_g^X attached to $g \in G^X$ is given by the order, denoted n_g, of the image of g in $\bar{G}^X$. (Cf. §A.3 for the values n_g.)

The Niemeier root systems and their corresponding umbral groups are described in Table 1. The root systems are given in terms of their simple components of ADE type. Here $D_{10}E_7^2$, for example, means the direct sum of one copy of the D_{10} root system and two copies of the E_7 root system. The symbol ℓ is called the *lambency* of X, and the Coxeter number m^X appears as the first summand of ℓ.

In the descriptions of the umbral groups G^X, and their permutation group quotients $\bar{G}^X$, we write M_{24} and M_{12} for the sporadic simple groups of Mathieu which act quintuply transitively on 24 and 12 points, respectively. (Cf. e.g. [21].) We write $GL_n(q)$ for the general linear group of a vector space of dimension n over a field with q elements, and $SL_n(q)$ is the subgroup of linear transformations with determinant 1, &c. The symbols $AGL_3(2)$ denote the *affine general linear group*, obtained by adjoining translations to $GL_3(2)$. We write Dih_n for the dihedral group of order $2n$, and Sym_n denotes the symmetric group on n symbols. We use n as a shorthand for a cyclic group of order n.

Table 1: The Umbral Groups

X	A_1^{24}	A_2^{12}	A_3^8	A_4^6	$A_5^4D_4$	A_6^4	$A_7^2D_5^2$
ℓ	2	3	4	5	6	7	8
G^X	M_{24}	$2.M_{12}$	$2.AGL_3(2)$	$GL_2(5)/2$	$GL_2(3)$	$SL_2(3)$	Dih_4
$\bar{G}^X$	M_{24}	M_{12}	$AGL_3(2)$	$PGL_2(5)$	$PGL_2(3)$	$PSL_2(3)$	2^2

X	A_8^3	$A_9^2D_6$	$A_{11}D_7E_6$	A_{12}^2	$A_{15}D_9$	$A_{17}E_7$	A_{24}
ℓ	9	10	12	13	16	18	25
G^X	Dih_6	4	2	4	2	2	2
$\bar{G}^X$	Sym_3	2	1	2	1	1	1

X	D_4^6	D_6^4	D_8^3	$D_{10}E_7^2$	D_{12}^2	$D_{16}E_8$	D_{24}
ℓ	6+3	10+5	14+7	18+9	22+11	30+15	46+23
G^X	$3.Sym_6$	Sym_4	Sym_3	2	2	1	1
$\bar{G}^X$	Sym_6	Sym_4	Sym_3	2	2	1	1

X	E_6^4	E_8^3
ℓ	12+4	30+6,10,15
G^X	$GL_2(3)$	Sym_3
$\bar{G}^X$	$PGL_2(3)$	Sym_3

We also use the notational convention of writing $A.B$ to denote the middle term in a short exact sequence $1 \to A \to A.B \to B \to 1$. This introduces some ambiguity which is nonetheless easily navigated in practice. For example, $2.M_{12}$ is the unique (up to isomorphism) double cover of M_{12} which is not $2 \times M_{12}$. The group $AGL_3(2)$ naturally embeds in $GL_4(2)$, which in turn admits a unique (up to isomorphism) double cover $2.GL_4(2)$ which is not a direct product. The group we denote $2.AGL_3(2)$ is the preimage of $AGL_3(2) < GL_4(2)$ in $2.GL_4(2)$ under the natural projection.

A.2 Character Tables

Table 2: Character table of $G^X \simeq M_{24}$, $X = A_1^{24}$

$[g]$	FS	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	23A	23B
$[g^2]$		1A	1A	1A	3A	3B	2A	2A	2B	5A	3A	3B	7A	7B	4B	5A	11A	6A	6B	7A	7B	15A	15B	21A	21B	23A	23B
$[g^3]$		1A	2A	2B	1A	1A	4A	4B	4C	5A	2A	2B	7B	7A	8A	10A	11A	4A	4C	14B	14A	5A	5A	7B	7A	23A	23B
$[g^5]$		1A	2A	2B	3A	3B	4A	4B	4C	1A	6A	6B	7B	7A	8A	2B	11A	12A	12B	14B	14A	3A	3A	21B	21A	23B	23A
$[g^7]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	1A	1A	8A	10A	11A	12A	12B	2A	2A	15B	15A	3B	3B	23B	23A
$[g^{11}]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	1A	12A	12B	14A	14B	15B	15A	21A	21B	23B	23A
$[g^{23}]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	1A	1A
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	+	23	7	−1	5	−1	−1	3	−1	3	1	−1	2	2	1	−1	1	−1	−1	0	0	0	0	−1	−1	0	0
χ_3	○	45	−3	5	0	3	−3	1	1	0	0	−1	b_7	$\overline{b_7}$	−1	0	1	0	1	$-b_7$	$-\overline{b_7}$	0	0	b_7	$\overline{b_7}$	−1	−1
χ_4	○	45	−3	5	0	3	−3	1	1	0	0	−1	$\overline{b_7}$	b_7	−1	0	1	0	1	$-\overline{b_7}$	$-b_7$	0	0	$\overline{b_7}$	b_7	−1	−1
χ_5	○	231	7	−9	−3	0	−1	−1	3	1	1	0	0	0	−1	1	0	−1	0	0	0	b_{15}	$\overline{b_{15}}$	0	0	1	1
χ_6	○	231	7	−9	−3	0	−1	−1	3	1	1	0	0	0	−1	1	0	−1	0	0	0	$\overline{b_{15}}$	b_{15}	0	0	1	1
χ_7	+	252	28	12	9	0	4	4	0	2	1	0	0	0	0	2	−1	1	0	0	0	−1	−1	0	0	−1	−1
χ_8	+	253	13	−11	10	1	−3	1	1	3	−2	1	1	1	−1	−1	0	0	1	−1	−1	0	0	1	1	0	0
χ_9	+	483	35	3	6	0	3	3	3	−2	2	0	0	0	−1	−2	−1	0	0	0	0	1	1	0	0	0	0
χ_{10}	○	770	−14	10	5	−7	2	−2	−2	0	1	1	0	0	0	0	0	−1	1	0	0	0	0	0	0	b_{23}	$\overline{b_{23}}$
χ_{11}	○	770	−14	10	5	−7	2	−2	−2	0	1	1	0	0	0	0	0	−1	1	0	0	0	0	0	0	$\overline{b_{23}}$	b_{23}
χ_{12}	○	990	−18	−10	0	3	6	2	−2	0	0	−1	b_7	$\overline{b_7}$	0	0	0	0	1	b_7	$\overline{b_7}$	0	0	b_7	$\overline{b_7}$	1	1
χ_{13}	○	990	−18	−10	0	3	6	2	−2	0	0	−1	$\overline{b_7}$	b_7	0	0	0	0	1	$\overline{b_7}$	b_7	0	0	$\overline{b_7}$	b_7	1	1
χ_{14}	+	1035	27	35	0	6	3	−1	3	0	0	2	−1	−1	1	0	1	0	0	−1	−1	0	0	−1	−1	0	0
χ_{15}	○	1035	−21	−5	0	−3	3	3	−1	0	0	1	$2b_7$	$2\overline{b_7}$	−1	0	1	0	−1	0	0	0	0	$-b_7$	$-\overline{b_7}$	0	0
χ_{16}	○	1035	−21	−5	0	−3	3	3	−1	0	0	1	$2\overline{b_7}$	$2b_7$	−1	0	1	0	−1	0	0	0	0	$-\overline{b_7}$	$-b_7$	0	0
χ_{17}	+	1265	49	−15	5	8	−7	1	−3	0	1	0	−2	−2	1	0	0	−1	0	0	0	0	0	1	1	0	0
χ_{18}	+	1771	−21	11	16	7	3	−5	−1	1	0	−1	0	0	−1	1	0	0	−1	0	0	1	1	0	0	0	0
χ_{19}	+	2024	8	24	−1	8	8	0	0	−1	−1	0	1	1	0	−1	0	−1	0	1	1	−1	−1	1	1	0	0
χ_{20}	+	2277	21	−19	0	6	−3	1	−3	−3	0	2	2	2	−1	1	0	0	0	0	0	0	0	−1	−1	0	0
χ_{21}	+	3312	48	16	0	−6	0	0	0	−3	0	−2	1	1	0	1	1	0	0	−1	−1	0	0	1	1	0	0
χ_{22}	+	3520	64	0	10	−8	0	0	0	0	−2	0	−1	−1	0	0	0	0	0	1	1	0	0	−1	−1	1	1
χ_{23}	+	5313	49	9	−15	0	1	−3	−3	3	1	0	0	0	−1	−1	0	1	0	0	0	0	0	0	0	0	0
χ_{24}	+	5544	−56	24	9	0	−8	0	0	−1	1	0	0	0	0	−1	0	1	0	0	0	−1	−1	0	0	1	1
χ_{25}	+	5796	−28	36	−9	0	−4	4	0	1	−1	0	0	0	0	1	−1	−1	0	0	0	1	1	0	0	0	0
χ_{26}	+	10395	−21	−45	0	0	3	−1	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	−1	−1

Table 3: Character table of $G^X \simeq 2.M_{12}$, $X = A_2^{12}$

$[g]$	FS	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8A	8B	8C	8D	20A	20B	11A	22A	11B	22B
$[g^2]$		1A	1A	2A	1A	1A	3A	3A	3B	3B	2B	2B	5A	5A	6B	3A	3A	4B	4B	4C	4C	10A	10A	11B	11B	11A	11A
$[g^3]$		1A	2A	4A	2B	2C	1A	2A	1A	2A	4B	4C	5A	10A	4A	2B	2C	8A	8B	8C	8D	20A	20B	11A	22A	11B	22B
$[g^5]$		1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	1A	2A	12A	6C	6D	8B	8A	8D	8C	4A	4A	11A	22A	11B	22B
$[g^{11}]$		1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8A	8B	8C	8D	20B	20A	1A	2A	1A	2A
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	+	11	11	−1	3	3	2	2	−1	−1	−1	3	1	1	−1	0	0	−1	−1	1	1	−1	−1	0	0	0	0
χ_3	+	11	11	−1	3	3	2	2	−1	−1	3	−1	1	1	−1	0	0	1	1	−1	−1	−1	−1	0	0	0	0
χ_4	○	16	16	4	0	0	−2	−2	1	1	0	0	1	1	1	0	0	0	0	0	0	−1	−1	b_{11}	b_{11}	$\overline{b_{11}}$	$\overline{b_{11}}$
χ_5	○	16	16	4	0	0	−2	−2	1	1	0	0	1	1	1	0	0	0	0	0	0	−1	−1	$\overline{b_{11}}$	$\overline{b_{11}}$	b_{11}	b_{11}
χ_6	+	45	45	5	−3	−3	0	0	3	3	1	1	0	0	−1	0	0	−1	−1	−1	−1	0	0	1	1	1	1
χ_7	+	54	54	6	6	6	0	0	0	0	2	2	−1	−1	0	0	0	0	0	0	0	1	1	−1	−1	−1	−1
χ_8	+	55	55	−5	7	7	1	1	1	1	−1	−1	0	0	1	1	1	−1	−1	−1	−1	0	0	0	0	0	0
χ_9	+	55	55	−5	−1	−1	1	1	1	1	3	−1	0	0	1	−1	−1	−1	−1	1	1	0	0	0	0	0	0
χ_{10}	+	55	55	−5	−1	−1	1	1	1	1	−1	3	0	0	1	−1	−1	1	1	−1	−1	0	0	0	0	0	0
χ_{11}	+	66	66	6	2	2	3	3	0	0	−2	−2	1	1	0	−1	−1	0	0	0	0	1	1	0	0	0	0
χ_{12}	+	99	99	−1	3	3	0	0	3	3	−1	−1	−1	−1	−1	0	0	1	1	1	1	−1	−1	0	0	0	0
χ_{13}	+	120	120	0	−8	−8	3	3	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	−1	−1	−1	−1
χ_{14}	+	144	144	4	0	0	0	0	−3	−3	0	0	−1	−1	1	0	0	0	0	0	0	−1	−1	1	1	1	1
χ_{15}	+	176	176	−4	0	0	−4	−4	−1	−1	0	0	1	1	−1	0	0	0	0	0	0	1	1	0	0	0	0
χ_{16}	○	10	−10	0	−2	2	1	−1	−2	2	0	0	0	0	0	1	−1	a_2	$\overline{a_2}$	a_2	$\overline{a_2}$	0	0	−1	1	−1	1
χ_{17}	○	10	−10	0	−2	2	1	−1	−2	2	0	0	0	0	0	1	−1	$\overline{a_2}$	a_2	$\overline{a_2}$	a_2	0	0	−1	1	−1	1
χ_{18}	+	12	−12	0	4	−4	3	−3	0	0	0	0	2	−2	0	1	−1	0	0	0	0	0	0	1	−1	1	−1
χ_{19}	−	32	−32	0	0	0	−4	4	2	−2	0	0	2	−2	0	0	0	0	0	0	0	0	0	−1	1	−1	1
χ_{20}	○	44	−44	0	4	−4	−1	1	2	−2	0	0	−1	1	0	1	−1	0	0	0	0	a_5	$\overline{a_5}$	0	0	0	0
χ_{21}	○	44	−44	0	4	−4	−1	1	2	−2	0	0	−1	1	0	1	−1	0	0	0	0	$\overline{a_5}$	a_5	0	0	0	0
χ_{22}	○	110	−110	0	−6	6	2	−2	2	−2	0	0	0	0	0	0	0	a_2	$\overline{a_2}$	$\overline{a_2}$	a_2	0	0	0	0	0	0
χ_{23}	○	110	−110	0	−6	6	2	−2	2	−2	0	0	0	0	0	0	0	$\overline{a_2}$	a_2	a_2	$\overline{a_2}$	0	0	0	0	0	0
χ_{24}	+	120	−120	0	8	−8	3	−3	0	0	0	0	0	0	0	−1	1	0	0	0	0	0	0	−1	1	−1	1
χ_{25}	○	160	−160	0	0	0	−2	2	−2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	$-b_{11}$	b_{11}	$-\overline{b_{11}}$	$\overline{b_{11}}$
χ_{26}	○	160	−160	0	0	0	−2	2	−2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	$-\overline{b_{11}}$	$\overline{b_{11}}$	$-b_{11}$	b_{11}

Table 4: Character table of $G^X \simeq 2.AGL_3(2)$, $X = A_3^8$

$[g]$	FS	1A	2A	2B	4A	4B	2C	3A	6A	6B	6C	8A	4C	7A	14A	7B	14B
$[g^2]$		1A	1A	1A	2A	2B	1A	3A	3A	3A	3A	4A	2C	7A	7A	7B	7B
$[g^3]$		1A	2A	2B	4A	4B	2C	1A	2A	2B	2B	8A	4C	7B	14B	7A	14A
$[g^7]$		1A	2A	2B	4A	4B	2C	3A	6A	6B	6C	8A	4C	1A	2A	1A	2A
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	○	3	3	3	-1	-1	-1	0	0	0	0	1	1	b_7	b_7	$\overline{b_7}$	$\overline{b_7}$
χ_3	○	3	3	3	-1	-1	-1	0	0	0	0	1	1	$\overline{b_7}$	$\overline{b_7}$	b_7	b_7
χ_4	+	6	6	6	2	2	2	0	0	0	0	0	0	-1	-1	-1	-1
χ_5	+	7	7	7	-1	-1	-1	1	1	1	1	-1	-1	0	0	0	0
χ_6	+	8	8	8	0	0	0	-1	-1	-1	-1	0	0	1	1	1	1
χ_7	+	7	7	-1	3	-1	-1	1	1	-1	-1	1	-1	0	0	0	0
χ_8	+	7	7	-1	-1	-1	3	1	1	-1	-1	-1	1	0	0	0	0
χ_9	+	14	14	-2	2	-2	2	-1	-1	1	1	0	0	0	0	0	0
χ_{10}	+	21	21	-3	1	1	-3	0	0	0	0	-1	1	0	0	0	0
χ_{11}	+	21	21	-3	-3	1	1	0	0	0	0	1	-1	0	0	0	0
χ_{12}	+	8	-8	0	0	0	0	2	-2	0	0	0	0	1	-1	1	-1
χ_{13}	○	8	-8	0	0	0	0	-1	1	a_3	$\overline{a_3}$	0	0	1	-1	1	-1
χ_{14}	○	8	-8	0	0	0	0	-1	1	$\overline{a_3}$	a_3	0	0	1	-1	1	-1
χ_{15}	○	24	-24	0	0	0	0	0	0	0	0	0	0	$\overline{b_7}$	$-\overline{b_7}$	b_7	$-b_7$
χ_{16}	○	24	-24	0	0	0	0	0	0	0	0	0	0	b_7	$-b_7$	$\overline{b_7}$	$-\overline{b_7}$

Table 5: Character table of $G^X \simeq GL_2(5)/2$, $X = A_4^6$

$[g]$	FS	1A	2A	2B	2C	3A	6A	5A	10A	4A	4B	4C	4D	12A	12B
$[g^2]$		1A	1A	1A	1A	3A	3A	5A	5A	2A	2A	2C	2C	6A	6A
$[g^3]$		1A	2A	2B	2C	1A	2A	5A	10A	4B	4A	4D	4C	4B	4A
$[g^5]$		1A	2A	2B	2C	3A	6A	1A	2A	4A	4B	4C	4D	12A	12B
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	+	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1
χ_3	+	4	4	0	0	1	1	-1	-1	2	2	0	0	-1	-1
χ_4	+	4	4	0	0	1	1	-1	-1	-2	-2	0	0	1	1
χ_5	+	5	5	1	1	-1	-1	0	0	1	1	-1	-1	1	1
χ_6	+	5	5	1	1	-1	-1	0	0	-1	-1	1	1	-1	-1
χ_7	+	6	6	-2	-2	0	0	1	1	0	0	0	0	0	0
χ_8	○	1	-1	1	-1	1	-1	1	-1	a_1	$-a_1$	a_1	$-a_1$	a_1	$-a_1$
χ_9	○	1	-1	1	-1	1	-1	1	-1	$-a_1$	a_1	$-a_1$	a_1	$-a_1$	a_1
χ_{10}	○	4	-4	0	0	1	-1	-1	1	$2a_1$	$-2a_1$	0	0	$-a_1$	a_1
χ_{11}	○	4	-4	0	0	1	-1	-1	1	$-2a_1$	$2a_1$	0	0	a_1	$-a_1$
χ_{12}	○	5	-5	1	-1	-1	1	0	0	a_1	$-a_1$	$-a_1$	a_1	a_1	$-a_1$
χ_{13}	○	5	-5	1	-1	-1	1	0	0	$-a_1$	a_1	a_1	$-a_1$	$-a_1$	a_1
χ_{14}	+	6	-6	-2	2	0	0	1	-1	0	0	0	0	0	0

Table 6: Character table of $G^X \simeq GL_2(3)$, $X \in \{A_5^4D_4, E_6^4\}$

$[g]$	FS	1A	2A	2B	4A	3A	6A	8A	8B
$[g^2]$		1A	1A	1A	2A	3A	3A	4A	4A
$[g^3]$		1A	2A	2B	4A	1A	2A	8A	8B
χ_1	$+$	1	1	1	1	1	1	1	1
χ_2	$+$	1	1	-1	1	1	1	-1	-1
χ_3	$+$	2	2	0	2	-1	-1	0	0
χ_4	$+$	3	3	-1	-1	0	0	1	1
χ_5	$+$	3	3	1	-1	0	0	-1	-1
χ_6	$\circ$	2	-2	0	0	-1	1	a_2	$\overline{a_2}$
χ_7	$\circ$	2	-2	0	0	-1	1	$\overline{a_2}$	a_2
χ_8	$+$	4	-4	0	0	1	-1	0	0

Table 7: Character table of $G^X \simeq 3.Sym_6$, $X = D_4^6$

$[g]$	FS	1A	3A	2A	6A	3B	3C	4A	12A	5A	15A	15B	2B	2C	4B	6B	6C
$[g^2]$		1A	3A	1A	3A	3B	3C	2A	6A	5A	15A	15B	1A	1A	2A	3B	3C
$[g^3]$		1A	1A	2A	2A	1A	1A	4A	4A	5A	5A	5A	2B	2C	4B	2B	2C
$[g^5]$		1A	3A	2A	6A	3B	3C	4A	12A	1A	3A	3A	2B	2C	4B	6B	6C
χ_1	$+$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	$+$	1	1	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1
χ_3	$+$	5	5	1	1	2	-1	-1	-1	0	0	0	3	-1	1	0	-1
χ_4	$+$	5	5	1	1	2	-1	-1	-1	0	0	0	-3	1	-1	0	1
χ_5	$+$	5	5	1	1	-1	2	-1	-1	0	0	0	-1	3	1	-1	0
χ_6	$+$	5	5	1	1	-1	2	-1	-1	0	0	0	1	-3	-1	1	0
χ_7	$+$	16	16	0	0	-2	-2	0	0	1	1	1	0	0	0	0	0
χ_8	$+$	9	9	1	1	0	0	1	1	-1	-1	-1	3	3	-1	0	0
χ_9	$+$	9	9	1	1	0	0	1	1	-1	-1	-1	-3	-3	1	0	0
χ_{10}	$+$	10	10	-2	-2	1	1	0	0	0	0	0	2	-2	0	-1	1
χ_{11}	$+$	10	10	-2	-2	1	1	0	0	0	0	0	-2	2	0	1	-1
χ_{12}	$\circ$	6	-3	-2	1	0	0	2	-1	1	b_{15}	$\overline{b_{15}}$	0	0	0	0	0
χ_{13}	$\circ$	6	-3	-2	1	0	0	2	-1	1	$\overline{b_{15}}$	b_{15}	0	0	0	0	0
χ_{14}	$+$	12	-6	4	-2	0	0	0	0	2	-1	-1	0	0	0	0	0
χ_{15}	$+$	18	-9	2	-1	0	0	2	-1	-2	1	1	0	0	0	0	0
χ_{16}	$+$	30	-15	-2	1	0	0	-2	1	0	0	0	0	0	0	0	0

Table 8: Character table of $G^X \simeq SL_2(3)$, $X = A_6^4$

$[g]$	FS	1A	2A	4A	3A	6A	3B	6B
$[g^2]$		1A	1A	2A	3B	3A	3A	3B
$[g^3]$		1A	2A	4A	1A	2A	1A	2A
χ_1	$+$	1	1	1	1	1	1	1
χ_2	$\circ$	1	1	1	b_3	$\overline{b_3}$	$\overline{b_3}$	b_3
χ_3	$\circ$	1	1	1	$\overline{b_3}$	b_3	b_3	$\overline{b_3}$
χ_4	$+$	3	3	-1	0	0	0	0
χ_5	$-$	2	-2	0	-1	1	-1	1
χ_6	$\circ$	2	-2	0	$-\overline{b_3}$	b_3	$-b_3$	$\overline{b_3}$
χ_7	$\circ$	2	-2	0	$-b_3$	$\overline{b_3}$	$-\overline{b_3}$	b_3

Table 9: Character table of $G^X \simeq Dih_4$, $X = A_7^2 D_5^2$

$[g]$	FS	1A	2A	2B	2C	4A
$[g^2]$		1A	1A	1A	1A	2A
χ_1	$+$	1	1	1	1	1
χ_2	$+$	1	1	-1	-1	1
χ_3	$+$	1	1	-1	1	-1
χ_4	$+$	1	1	1	-1	-1
χ_5	$+$	2	-2	0	0	0

Table 10: Character table of $G^X \simeq Dih_6$, $X = A_8^3$

$[g]$	FS	1A	2A	2B	2C	3A	6A
$[g^2]$		1A	1A	1A	1A	3A	3A
$[g^3]$		1A	2A	2B	2C	1A	2A
χ_1	$+$	1	1	1	1	1	1
χ_2	$+$	1	1	-1	-1	1	1
χ_3	$+$	2	2	0	0	-1	-1
χ_4	$+$	1	-1	-1	1	1	-1
χ_5	$+$	1	-1	1	-1	1	-1
χ_6	$+$	2	-2	0	0	-1	1

Table 11: Character table of $G^X \simeq 4$, for $X \in \{A_9^2D_6, A_{12}^2\}$

$[g]$	FS	1A	2A	4A	4B
$[g^2]$		1A	1A	2A	2A
χ_1	$+$	1	1	1	1
χ_2	$+$	1	1	-1	-1
χ_3	$\circ$	1	-1	a_1	$\overline{a_1}$
χ_4	$\circ$	1	-1	$\overline{a_1}$	a_1

Table 12: Character table of $G^X \simeq PGL_2(3) \simeq Sym_4$, $X = D_6^4$

$[g]$	FS	1A	2A	3A	2B	4A
$[g^2]$		1A	1A	3A	1A	2A
$[g^3]$		1A	2A	1A	2B	4A
χ_1	$+$	1	1	1	1	1
χ_2	$+$	1	1	1	-1	-1
χ_3	$+$	2	2	-1	0	0
χ_4	$+$	3	-1	0	1	-1
χ_5	$+$	3	-1	0	-1	1

Table 13: Character table of $G^X \simeq 2$, for $X \in \{A_{11}D_7E_6, A_{15}D_9, A_{17}E_7, A_{24}, D_{10}E_7^2, D_{12}^2\}$

$[g]$	FS	1A	2A
$[g^2]$		1A	1A
χ_1	$+$	1	1
χ_2	$+$	1	-1

Table 14: Character table of $G^X \simeq Sym_3$, $X \in \{D_8^3, E_8^3\}$

$[g]$	FS	1A	2A	3A
$[g^2]$		1A	1A	3A
$[g^3]$		1A	2A	1A
χ_1	$+$	1	1	1
χ_2	$+$	1	-1	1
χ_3	$+$	2	0	-1

A.3 Twisted Euler Characters

In this section we reproduce certain characters—the *twisted Euler characters*—which are attached to each group G^X, via its action on the root system X. (Their construction is described in detail in §2.4 of [16].)

To interpret the tables, write X_A for the (possibly empty) union of type A components of X, and interpret X_D and X_E similarly, so that if $m = m^X$ Then $X = A_{m-1}^d$ for some d, and $X = X_A \cup X_D \cup X_E$, for example. Then $g \mapsto \bar{\chi}_g^{X_A}$ denotes the character of the permutation representation attached to the action of $\bar{G}^X$ on the simple components of X_A. The characters $g \mapsto \bar{\chi}_g^{X_D}$ and $g \mapsto \bar{\chi}_g^{X_E}$ are defined similarly. The characters $\chi_g^{X_A}$, $\chi_g^{X_D}$, $\chi_g^{X_E}$ and $\check{\chi}_g^{X_D}$ incorporate outer automorphisms of simple root systems induced by the action G^X on X. We refer to §2.4 of [16] for full details of the construction. For the purposes of this work, it suffices to have the explicit descriptions in the tables in this section. The twisted Euler characters presented here will be used to specify the umbral shadow functions in §B.2.

The twisted Euler character tables also attach integers n_g and h_g to each $g \in G^X$. By definition, n_g is the order of the image of $g \in G^X$ in $\bar{G}^X$ (cf. §A.1). The integer h_g may be defined by setting $h_g := N_g/n_g$ where N_g is the product of the shortest and longest cycle lengths appearing in the cycle shape attached to g by the action of G^X on a (suitable) set of simple roots for X.

Table 15: Twisted Euler characters at $\ell = 2$, $X = A_1^{24}$

$[g]$	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7AB	8A	10A	11A	12A	12B	14AB	15AB	21AB	23AB
$n_g\|h_g$	1|1	2|1	2|2	3|1	3|3	4|2	4|1	4|4	5|1	6|1	6|6	7|1	8|1	10|2	11|1	12|2	12|12	14|1	15|1	21|3	23|1
$\bar{\chi}_g^{X_A}$	24	8	0	6	0	0	4	0	4	2	0	3	2	0	2	0	0	1	1	0	1

Table 16: Twisted Euler characters at $\ell = 3$, $X = A_2^{12}$

$[g]$	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8AB	8CD	20AB	11AB	22AB
$n_g\|h_g$	1|1	1|4	2|8	2|1	2|2	3|1	3|4	3|3	3|12	4|2	4|1	5|1	5|4	6|24	6|1	6|2	8|4	8|1	10|8	11|1	11|4
$\bar{\chi}_g^{X_A}$	12	12	0	4	4	3	3	0	0	0	4	2	2	0	1	1	0	2	0	1	1
$\chi_g^{X_A}$	12	−12	0	4	−4	3	−3	0	0	0	0	2	−2	0	1	−1	0	0	0	1	−1

Table 17: Twisted Euler characters at $\ell = 4$, $X = A_3^8$

$[g]$	1A	2A	2B	4A	4B	2C	3A	6A	6BC	8A	4C	7AB	14AB
$n_g\|h_g$	1|1	1|2	2|2	2|4	4|4	2|1	3|1	3|2	6|2	4|8	4|1	7|1	7|2
$\bar{\chi}_g^{X_A}$	8	8	0	0	0	4	2	2	0	0	2	1	1
$\chi_g^{X_A}$	8	−8	0	0	0	0	2	−2	0	0	0	1	−1

Table 18: Twisted Euler characters at $\ell = 5$, $X = A_4^6$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
$n_g\|h_g$	1|1	1|4	2|2	2|1	3|3	3|12	5|1	5|4	2|8	4|1	6|24
$\bar{\chi}_g^{X_A}$	6	6	2	2	0	0	1	1	0	2	0
$\chi_g^{X_A}$	6	−6	−2	2	0	0	1	−1	0	0	0

Table 19: Twisted Euler characters at $\ell = 6$, $X = A_5^4 D_4$

$[g]$	1A	2A	2B	4A	3A	6A	8AB
$n_g\|h_g$	1|1	1|2	2|1	2|2	3|1	3|2	4|2
$\bar{\chi}_g^{X_A}$	4	4	2	0	1	1	0
$\chi_g^{X_A}$	4	−4	0	0	1	−1	0
$\bar{\chi}_g^{X_D}$	1	1	1	1	1	1	1
$\chi_g^{X_D}$	1	1	−1	1	1	1	−1
$\check{\chi}_g^{X_D}$	2	2	0	2	−1	−1	0

Table 20: Twisted Euler characters at $\ell = 6 + 3$, $X = D_4^6$

$[g]$	1A	3A	2A	6A	3B	6C	4A	12A	5A	15AB	2B	2C	4B	6B	6C
$n_g\|h_g$	1|1	1|3	2|1	2|3	3|1	3|3	4|2	4|6	5|1	5|3	2|1	2|2	4|1	6|1	6|6
$\bar{\chi}_g^{X_D}$	6	6	2	2	3	0	0	0	1	1	4	0	2	1	0
$\chi_g^{X_D}$	6	6	2	2	3	0	0	0	1	1	−4	0	−2	−1	0
$\check{\chi}_g^{X_D}$	12	−6	4	−2	0	0	0	0	2	−1	0	0	0	0	0

Table 21: Twisted Euler characters at $\ell = 7$, $X = A_6^4$

$[g]$	1A	2A	4A	3AB	6AB
$n_g\vert h_g$	$1\vert 1$	$1\vert 4$	$2\vert 8$	$3\vert 1$	$3\vert 4$
$\bar{\chi}_g^{X_A}$	4	4	0	1	1
$\chi_g^{X_A}$	4	-4	0	1	-1

Table 22: Twisted Euler characters at $\ell = 8$, $X = A_7^2 D_5^2$

$[g]$	1A	2A	2B	2C	4A
$n_g\vert h_g$	$1\vert 1$	$1\vert 2$	$2\vert 1$	$2\vert 1$	$2\vert 4$
$\bar{\chi}_g^{X_A}$	2	2	0	2	0
$\chi_g^{X_A}$	2	-2	0	0	0
$\bar{\chi}_g^{X_D}$	2	2	2	0	0
$\chi_g^{X_D}$	2	-2	0	0	0

Table 23: Twisted Euler characters at $\ell = 9$, $X = A_8^3$

$[g]$	1A	2A	2B	2C	3A	6A
$n_g\vert h_g$	$1\vert 1$	$1\vert 4$	$2\vert 1$	$2\vert 2$	$3\vert 3$	$3\vert 12$
$\bar{\chi}_g^{X_A}$	3	3	1	1	0	0
$\chi_g^{X_A}$	3	-3	1	-1	0	0

Table 24: Twisted Euler characters at $\ell = 10$, $X = A_9^2 D_6$

$[g]$	1A	2A	4AB
$n_g\vert h_g$	$1\vert 1$	$1\vert 2$	$2\vert 2$
$\bar{\chi}_g^{X_A}$	2	2	0
$\chi_g^{X_A}$	2	-2	0
$\bar{\chi}_g^{X_D}$	1	1	1
$\chi_g^{X_D}$	1	1	-1

Table 25: Twisted Euler characters at $\ell = 10 + 5$, $X = D_6^4$

$[g]$	1A	2A	3A	2B	4A
$n_g\|h_g$	1|1	2|2	3|1	2|1	4|4
$\bar{\chi}_g^{X_D}$	4	0	1	2	0
$\chi_g^{X_D}$	4	0	1	-2	0

Table 26: Twisted Euler characters at $\ell = 12$, $X = A_{11}D_7E_6$

$[g]$	1A	2A
$n_g\|h_g$	1|1	1|2
$\bar{\chi}_g^{X_A}$	1	1
$\chi_g^{X_A}$	1	-1
$\bar{\chi}_g^{X_D}$	1	1
$\chi_g^{X_D}$	1	-1
$\bar{\chi}_g^{X_E}$	1	1
$\chi_g^{X_E}$	1	-1

Table 27: Twisted Euler characters at $\ell = 12 + 4$, $X = E_6^4$

$[g]$	1A	2A	2B	4A	3A	6A	8AB
$n_g\|h_g$	1|1	1|2	2|1	2|4	3|1	3|2	4|8
$\bar{\chi}_g^{X_E}$	4	4	2	0	1	1	0
$\chi_g^{X_E}$	4	-4	0	0	1	-1	0

Table 28: Twisted Euler characters at $\ell = 13$, $X = A_{12}^2$

$[g]$	1A	2A	4AB
$n_g\|h_g$	1|1	1|4	2|8
$\bar{\chi}_g^{X_A}$	2	2	0
$\chi_g^{X_A}$	2	-2	0

Table 29: Twisted Euler characters at $\ell = 14 + 7$, $X = D_8^3$

$[g]$	1A	2A	3A
$n_g \vert h_g$	1|1	2|1	3|3
$\bar{\chi}_g^{X_D}$	3	1	0
$\chi_g^{X_D}$	3	1	0

Table 30: Twisted Euler characters at $\ell = 16$, $X = A_{15}D_9$

$[g]$	1A	2A
$n_g \vert h_g$	1|1	1|2
$\bar{\chi}_g^{X_A}$	1	1
$\chi_g^{X_A}$	1	−1
$\bar{\chi}_g^{X_D}$	1	1
$\chi_g^{X_D}$	1	−1

Table 31: Twisted Euler characters at $\ell = 18$, $X = A_{17}E_7$

$[g]$	1A	2A
$n_g \vert h_g$	1|1	1|2
$\bar{\chi}_g^{X_A}$	1	1
$\chi_g^{X_A}$	1	−1
$\bar{\chi}_g^{X_E}$	1	1

Table 32: Twisted Euler characters at $\ell = 18 + 9$, $X = D_{10}E_7^2$

$[g]$	1A	2A
$n_g \vert h_g$	1|1	2|1
$\bar{\chi}_g^{X_D}$	1	1
$\chi_g^{X_D}$	1	−1
$\bar{\chi}_g^{X_E}$	2	0

Table 33: Twisted Euler characters at $\ell = 22+11$, $X = D_{12}^2$

$[g]$	1A	2A
$n_g \vert h_g$	$1 \vert 1$	$2 \vert 2$
$\bar{\chi}_g^{X_D}$	2	0
$\chi_g^{X_D}$	2	0

Table 34: Twisted Euler characters at $\ell = 25$, $X = A_{24}$

$[g]$	1A	2A
$n_g \vert h_g$	$1 \vert 1$	$1 \vert 4$
$\bar{\chi}_g^{X_A}$	1	1
$\chi_g^{X_A}$	1	-1

Table 35: Twisted Euler characters at $\ell = 30+6,10,15$, $X = E_8^3$

$[g]$	1A	2A	3A
$n_g \vert h_g$	$1 \vert 1$	$2 \vert 1$	$3 \vert 3$
$\bar{\chi}_g^{X_E}$	3	1	0

B The Umbral McKay-Thompson Series

In this section we describe the umbral McKay-Thompson series in complete detail. In particular, we present explicit formulas for all the McKay-Thompson series attached to elements of the umbral groups by umbral moonshine in §B.3. Most of these expressions appeared first in [15,16], but some appear for the first time in this work.

In order to facilitate explicit formulations we recall certain standard functions in §B.1. We then, using the twisted Euler characters of §A.3, explicitly describe the shadow functions of umbral moonshine in §B.2. The umbral McKay–Thompson series defined in §B.3 may also be described in terms of Rademacher sums, according to the results of [17]. We present this description in §B.4.

B.1 Special Functions

Throughout this section we assume $q := e^{2\pi i\tau}$, and $u := e^{2\pi iz}$, where $\tau, z \in \mathbb{C}$ with Im $\tau > 0$. The Dedekind eta function is $\eta(\tau) := q^{1/24}\prod_{n>0}(1-q^n)$, where . Write $\Lambda_M(\tau)$ for the function

$$\Lambda_M(\tau) := Mq\frac{\mathrm{d}}{\mathrm{d}q}\left(\log\frac{\eta(M\tau)}{\eta(\tau)}\right) = \frac{M(M-1)}{24} + M\sum_{k>0}\sum_{d|k} d\left(q^k - Mq^{Mk}\right),$$

which is a modular form of weight two for $\Gamma_0(N)$ if $M|N$.

Define the Jacobi theta function $\theta_1(\tau, z)$ by setting

$$\theta_1(\tau, z) := iq^{1/8}u^{-1/2}\sum_{n\in\mathbb{Z}}(-1)^n u^n q^{n(n-1)/2}. \tag{B.1}$$

According to the Jacobi triple product identity we have

$$\theta_1(\tau, z) = -iq^{1/8}u^{1/2}\prod_{n>0}(1-u^{-1}q^{n-1})(1-uq^n)(1-q^n). \tag{B.2}$$

The other Jacobi theta functions are

$$\begin{aligned}
\theta_2(\tau, z) &:= q^{1/8}u^{1/2}\prod_{n>0}(1+u^{-1}q^{n-1})(1+uq^n)(1-q^n),\\
\theta_3(\tau, z) &:= \prod_{n>0}(1+u^{-1}q^{n-1/2})(1+uq^{n-1/2})(1-q^n),\\
\theta_4(\tau, z) &:= \prod_{n>0}(1-u^{-1}q^{n-1/2})(1-uq^{n-1/2})(1-q^n).
\end{aligned} \tag{B.3}$$

Define $\Psi_{1,1}$ and $\Psi_{1,-1/2}$ by setting

$$\begin{aligned}
\Psi_{1,1}(\tau, z) &:= -i\frac{\theta_1(\tau, 2z)\eta(\tau)^3}{\theta_1(\tau, z)^2},\\
\Psi_{1,-1/2}(\tau, z) &:= -i\frac{\eta(\tau)^3}{\theta_1(\tau, z)}.
\end{aligned} \tag{B.4}$$

These are meromorphic Jacobi forms of weight one, with indexes 1 and $-1/2$, respectively. Here, the term meromorphic refers to the presence of simple poles in the functions $z \mapsto \Psi_{1,*}(\tau, z)$, for fixed $\tau \in \mathbb{H}$, at lattice points $z \in \mathbb{Z}\tau + \mathbb{Z}$. (Cf. §8 of [22].)

From §5 of [29] we recall the *index m theta functions*, for $m \in \mathbb{Z}$, defined by setting

$$\theta_{m,r}(\tau,z) := \sum_{k\in\mathbb{Z}} u^{2mk+r} q^{(2mk+r)^2/4m}, \tag{B.5}$$

where $r \in \mathbb{Z}$. Evidently, $\theta_{m,r}$ only depends on $r \mod 2m$. We set $S_{m,r}(\tau) := \frac{1}{2\pi i}\, \partial_z \theta_{m,r}(\tau,z)\big|_{z=0}$, so that

$$S_{m,r}(\tau) = \sum_{k\in\mathbb{Z}} (2mk+r) q^{(2mk+r)^2/4m}. \tag{B.6}$$

For a m a positive integer define

$$\mu_{m,0}(\tau,z) = \sum_{k\in\mathbb{Z}} u^{2km} q^{mk^2} \frac{uq^k+1}{uq^k-1} = \frac{u+1}{u-1} + O(q), \tag{B.7}$$

and observe that we recover $\Psi_{1,1}$ upon specializing (B.7) to $m = 1$. Observe also that

$$\mu_{m,0}(\tau,z+1/2) = \sum_{k\in\mathbb{Z}} u^{2km} q^{mk^2} \frac{uq^k-1}{uq^k+1} = \frac{u-1}{u+1} + O(q). \tag{B.8}$$

Define the even and odd parts of $\mu_{m,0}$ by setting

$$\mu^k_{m,0}(\tau,z) := \frac{1}{2}(\mu_{m,0}(\tau,z) + (-1)^k \mu_{m,0}(\tau,z+1/2)) \tag{B.9}$$

for $k \mod 2$.

For $m, r \in \mathbb{Z} + \frac{1}{2}$ with $m > 0$ define *half-integral index theta functions*

$$\theta_{m,r}(\tau,z) := \sum_{k\in\mathbb{Z}} e(mk + r/2) u^{2mk+r} q^{(2mk+r)^2/4m}, \tag{B.10}$$

and define also $S_{m,r}(\tau) := \frac{1}{2\pi i} \partial_z \theta_{m,r}(\tau,z)|_{z=0}$, so that

$$S_{m,r}(\tau) = \sum_{k\in\mathbb{Z}} e(mk + r/2)(2mk+r) q^{(2mk+r)^2/4m}. \tag{B.11}$$

As in the integral index case, $\theta_{m,r}$ depends only on $r \mod 2m$. We recover $-\theta_1$ upon specializing $\theta_{m,r}$ to $m = r = 1/2$.

For $m \in \mathbb{Z} + 1/2$, $m > 0$, define

$$\mu_{m,0}(\tau,z) := i \sum_{k\in\mathbb{Z}} (-1)^k u^{2mk+1/2} q^{mk^2+k/2} \frac{1}{1-uq^k} = \frac{-iu^{1/2}}{y-1} + O(q). \tag{B.12}$$

Given $\alpha \in \mathbb{Q}$ write $[\alpha]$ for the operator on q-series (in rational, possibility negative powers of q) that eliminates exponents not contained in $\mathbb{Z} + \alpha$, so that if $f = \sum_{\beta\in\mathbb{Q}} c(\beta) q^\beta$ then

$$[\alpha]f := \sum_{n\in\mathbb{Z}} c(n+\alpha) q^{n+\alpha} \tag{B.13}$$

B.2 Shadows

Let X be a Niemeier root system and let $m = m^X$ be the Coxeter number of X. For $g \in G^X$ we define the associated *shadow function* $S^X_g = (S^X_{g,r})$ by setting

$$S^X_g := S^{X_A}_g + S^{X_D}_g + S^{X_E}_g \tag{B.14}$$

where the $S^{X_A}_g$, &c., are defined in the following way, in terms of the twisted Euler characters $\chi^{X_A}_g$, &c. given in §A.3, and the unary theta series $S_{m,r}$ (cf. (B.6)).

Note that if $m = m^X$ then $S^X_{g,r} = S^X_{g,r+2m} = -S^X_{g,-r}$ for all $g \in G^X$, so we need specify the $S^{X_A}_{g,r}$, &c., only for $0 < r < m$.

If $X_A = \emptyset$ then $S^{X_A}_g := 0$. Otherwise, we define $S^{X_A}_{g,r}$ for $0 < r < m$ by setting

$$S^{X_A}_{g,r} := \begin{cases} \chi^{X_A}_g S_{m,r} & \text{if } r = 0 \mod 2, \\ \bar{\chi}^{X_A}_g S_{m,r} & \text{if } r = 1 \mod 2. \end{cases} \tag{B.15}$$

If $X_D = \emptyset$ then $S^{X_D}_g := 0$. If $X_D \neq \emptyset$ then m is even and $m \geq 6$. If $m = 6$ then set

$$S^{X_D}_{g,r} := \begin{cases} 0 & \text{if } r = 0 \mod 2, \\ \bar{\chi}^{X_D}_g S_{6,r} + \chi^{X_D}_g S_{6,6-r} & \text{if } r = 1,5 \mod 6, \\ \check{\chi}^{X_D}_g S_{6,r} & \text{if } r = 3 \mod 6. \end{cases} \tag{B.16}$$

If $m > 6$ and $m = 2 \mod 4$ then set

$$S^{X_D}_{g,r} := \begin{cases} 0 & \text{if } r = 0 \mod 2, \\ \bar{\chi}^{X_D}_g S_{m,r} + \chi^{X_D}_g S_{m,m-r} & \text{if } r = 1 \mod 2. \end{cases} \tag{B.17}$$

If $m > 6$ and $m = 0 \mod 4$ then set

$$S^{X_D}_{g,r} := \begin{cases} \chi^{X_D}_g S_{m,m-r} & \text{if } r = 0 \mod 2, \\ \bar{\chi}^{X_D}_g S_{m,r} & \text{if } r = 1 \mod 2. \end{cases} \tag{B.18}$$

If $X_E = \emptyset$ then $S^{X_E}_g := 0$. Otherwise, m is 12 or 18 or 30. In case $m = 12$ define $S^{X_E}_{g,r}$ for $0 < r < 12$ by setting

$$S^{X_E}_{g,r} = \begin{cases} \bar{\chi}^{X_E}_g (S_{12,1} + S_{12,7}) & \text{if } r \in \{1,7\}, \\ \bar{\chi}^{X_E}_g (S_{12,5} + S_{12,11}) & \text{if } r \in \{5,11\}, \\ \chi^{X_E}_g (S_{12,4} + S_{12,8}) & \text{if } r \in \{4,8\}, \\ 0 & \text{else.} \end{cases} \tag{B.19}$$

In case $m = 18$ define $S^{X_E}_{g,r}$ for $0 < r < 18$ by setting

$$S^{X_E}_{g,r} = \begin{cases} \bar{\chi}^{X_E}_g (S_{18,r} + S_{18,18-r}) & \text{if } r \in \{1,5,7,11,13,17\}, \\ \bar{\chi}^{X_E}_g S_{18,9} & \text{if } r \in \{3,15\}, \\ \bar{\chi}^{X_E}_g (S_{18,3} + S_{18,9} + S_{18,15}) & \text{if } r = 9, \\ 0 & \text{else.} \end{cases} \tag{B.20}$$

In case $m = 30$ define $S^{X_E}_{g,r}$ for $0 < r < 30$ by setting

$$S^{X_E}_{g,r} = \begin{cases} \bar{\chi}^{X_E}_g (S_{30,1} + S_{30,11} + S_{30,19} + S_{30,29}) & \text{if } r \in \{1, 11, 19, 29\}, \\ \bar{\chi}^{X_E}_g (S_{30,7} + S_{30,13} + S_{30,17} + S_{30,23}) & \text{if } r \in \{7, 13, 17, 23\}, \\ 0 & \text{else.} \end{cases} \tag{B.21}$$

B.3 Explicit Prescriptions

Here we give explicit expressions for all the umbral McKay-Thompson series H^X_g. Most of these appeared first in [15,16]. The expressions in §§B.3.3, B.3.4, B.3.7, B.3.14 are taken from [26]. The expressions in §§B.3.11, B.3.15, B.3.19, B.3.23 are taken from [14]. The expressions for H^X_g with $X = E_8^3$ appeared first in [25]. The expression for $H^{(6+3)}_{2B,1}$ in §B.3.6, and the expressions for $H^{(12+4)}_{4A,r}$ and $H^{(12+4)}_{8AB,r}$ in §B.3.13, appear here for the first time.

The labels for conjugacy classes in G^X are as in §A.2.

B.3.1 $\ell = 2$, $X = A_1^{24}$

We have $G^{(2)} = G^X \simeq M_{24}$ and $m^X = 2$. So for $g \in M_{24}$, the associated umbral McKay-Thompson series $H^{(2)}_g = (H^{(2)}_{g,r})$ is a 4-vector-valued function, with components indexed by $r \in \mathbb{Z}/4\mathbb{Z}$, satisfying $H^{(2)}_{g,r} = -H^{(2)}_{g,-r}$, and in particular, $H^{(2)}_{g,r} = 0$ for $r = 0 \mod 2$. So it suffices to specify the $H^{(2)}_{g,1}$ explicitly.

Define $H^{(2)}_g = (H^{(2)}_{g,r})$ for $g = e$ by requiring that

$$-2\Psi_{1,1}(\tau, z)\varphi^{(2)}_1(\tau, z) = -24\mu_{2,0}(\tau, z) + \sum_{r \mod 4} H^{(2)}_{e,r}(\tau)\theta_{2,r}(\tau, z), \tag{B.22}$$

where

$$\varphi^{(2)}_1(\tau, z) := 4\left(\frac{\theta_2(\tau, z)^2}{\theta_2(\tau, 0)^2} + \frac{\theta_3(\tau, z)^2}{\theta_3(\tau, 0)^2} + \frac{\theta_4(\tau, z)^2}{\theta_4(\tau, 0)^2}\right). \tag{B.23}$$

More generally, for $g \in G^{(2)}$ define

$$H^{(2)}_{g,1}(\tau) := \frac{\bar{\chi}^{(2)}_g}{24} H^{(2)}_{e,1}(\tau) - F^{(2)}_g(\tau)\frac{1}{S_{2,1}(\tau)}, \tag{B.24}$$

where $\bar{\chi}^{(2)}_g$ and $F^{(2)}_g$ are as specified in Table 36. Note that $\bar{\chi}^{(2)}_g = \bar{\chi}^{X_A}_g$, the latter appearing in Table 15. Also, $S_{2,1}(\tau) = \eta(\tau)^3$.

The functions $f_{23,a}$ and $f_{23,b}$ in Table 36 are cusp forms of weight two for $\Gamma_0(23)$, defined by

$$\begin{aligned} f_{23,a}(\tau) &:= \frac{\eta(\tau)^3\eta(23\tau)^3}{\eta(2\tau)\eta(46\tau)} + 3\eta(\tau)^2\eta(23\tau)^2 + 4\eta(\tau)\eta(2\tau)\eta(23\tau)\eta(46\tau) + 4\eta(2\tau)^2\eta(46\tau)^2, \\ f_{23,b}(\tau) &:= \eta(\tau)^2\eta(23\tau)^2. \end{aligned} \tag{B.25}$$

Note that the definition of $F^{(2)}_g$ appearing here for $g \in 23A \cup 23B$ corrects errors in [11,12].

Table 36: Character Values and Weight Two Forms for $\ell = 2$, $X = A_1^{24}$

$[g]$	$\bar{\chi}_g^{(2)}$	$F_g^{(2)}(\tau)$
$1A$	24	0
$2A$	8	$16\Lambda_2(\tau)$
$2B$	0	$2\eta(\tau)^8\eta(2\tau)^{-4}$
$3A$	6	$6\Lambda_3(\tau)$
$3B$	0	$2\eta(\tau)^6\eta(3\tau)^{-2}$
$4A$	0	$2\eta(2\tau)^8\eta(4\tau)^{-4}$
$4B$	4	$4(-\Lambda_2(\tau)+\Lambda_4(\tau))$
$4C$	0	$2\eta(\tau)^4\eta(2\tau)^2\eta(4\tau)^{-2}$
$5A$	4	$2\Lambda_5(\tau)$
$6A$	2	$2(-\Lambda_2(\tau)-\Lambda_3(\tau)+\Lambda_6(\tau))$
$6B$	0	$2\eta(\tau)^2\eta(2\tau)^2\eta(3\tau)^2\eta(6\tau)^{-2}$
$7AB$	3	$\Lambda_7(\tau)$
$8A$	2	$-\Lambda_4(\tau)+\Lambda_8(\tau)$
$10A$	0	$2\eta(\tau)^3\eta(2\tau)\eta(5\tau)\eta(10\tau)^{-1}$
$11A$	2	$2(\Lambda_{11}(\tau)-11\eta(\tau)^2\eta(11\tau)^2)/5$
$12A$	0	$2\eta(\tau)^3\eta(4\tau)^2\eta(6\tau)^3\eta(2\tau)^{-1}\eta(3\tau)^{-1}\eta(12\tau)^{-2}$
$12B$	0	$2\eta(\tau)^4\eta(4\tau)\eta(6\tau)\eta(2\tau)^{-1}\eta(12\tau)^{-1}$
$14AB$	1	$(-\Lambda_2(\tau)-\Lambda_7(\tau)+\Lambda_{14}(\tau)-14\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau))/3$
$15AB$	1	$(-\Lambda_3(\tau)-\Lambda_5(\tau)+\Lambda_{15}(\tau)-15\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau))/4$
$21AB$	0	$(7\eta(\tau)^3\eta(7\tau)^3\eta(3\tau)^{-1}\eta(21\tau)^{-1}-\eta(\tau)^6\eta(3\tau)^{-2})/3$
$23AB$	1	$(\Lambda_{23}(\tau)-23f_{23,a}(\tau)-69f_{23,b}(\tau))/11$

B.3.2 $\ell = 3$, $X = A_2^{12}$

We have $G^{(3)} = G^X \simeq 2.M_{12}$ and $m^X = 3$. So for $g \in 2.M_{12}$, the associated umbral McKay-Thompson series $H_g^{(3)} = (H_{g,r}^{(3)})$ is a 6-vector-valued function, with components indexed by $r \in \mathbb{Z}/6\mathbb{Z}$, satisfying $H_{g,r}^{(3)} = -H_{g,-r}^{(3)}$, and in particular, $H_{g,r}^{(3)} = 0$ for $r = 0 \mod 3$. So it suffices to specify the $H_{g,1}^{(3)}$ and $H_{g,2}^{(3)}$ explicitly.

Define $H_g^{(3)} = (H_{g,r}^{(3)})$ for $g = e$ by requiring that

$$-2\Psi_{1,1}(\tau,z)\varphi_1^{(3)}(\tau,z) = -12\mu_{3,0}(\tau,z) + \sum_{r \mod 6} H_{e,r}^{(3)}(\tau)\theta_{3,r}(\tau,z), \tag{B.26}$$

where

$$\varphi_1^{(3)}(\tau,z) := 2\left(\frac{\theta_3(\tau,z)^2}{\theta_3(\tau,0)^2}\frac{\theta_4(\tau,z)^2}{\theta_4(\tau,0)^2} + \frac{\theta_4(\tau,z)^2}{\theta_4(\tau,0)^2}\frac{\theta_2(\tau,z)^2}{\theta_2(\tau,0)^2} + \frac{\theta_2(\tau,z)^2}{\theta_2(\tau,0)^2}\frac{\theta_3(\tau,z)^2}{\theta_3(\tau,0)^2}\right). \tag{B.27}$$

More generally, for $g \in G^{(3)}$ define

$$H^{(3)}_{g,1}(\tau) := \frac{\bar{\chi}^{(3)}_g}{12} H^{(3)}_{e,1}(\tau) + \frac{1}{2}\left(F^{(3)}_g + F^{(3)}_{zg}\right)\frac{1}{S_{3,1}(\tau)}, \tag{B.28}$$

$$H^{(3)}_{g,2}(\tau) := \frac{\chi^{(3)}_g}{12} H^{(3)}_{e,1}(\tau) + \frac{1}{2}\left(F^{(3)}_g - F^{(3)}_{zg}\right)\frac{1}{S_{3,2}(\tau)}, \tag{B.29}$$

where $\chi^{(3)}_g$ and $F^{(3)}_g$ are as specified in Table 37, and z is the non-trivial central element of $G^{(3)}$. The action of $g \mapsto zg$ on conjugacy classes can be read off Table 37, for the horizontal lines indicate the sets $[g] \cup [zg]$.

Note the eta product identities, $S_{3,1}(\tau) = \eta(2\tau)^5/\eta(4\tau)^2$, and $S_{3,2}(\tau) = 2\eta(\tau)^2\eta(4\tau)^2/\eta(2\tau)$. Note also that $\bar{\chi}^{(3)}_g = \bar{\chi}^{X_A}_g$ and $\chi^{(3)}_g = \chi^{X_A}_g$, the latter appearing in Table 16.

Table 37: Character Values and Weight Two Forms for $\ell = 3$, $X = A_2^{12}$

$[g]$	$\bar{\chi}^{(3)}_g$	$\chi^{(3)}_g$	$F^{(3)}_g(\tau)$
1A	12	12	0
2A	12	-12	0
4A	0	0	$-2\eta(\tau)^4\eta(2\tau)^2/\eta(4\tau)^2$
2B	4	4	$-16\Lambda_2(\tau)$
2C	4	-4	$16\Lambda_2(\tau) - \frac{16}{3}\Lambda_4(\tau)$
3A	3	3	$-6\Lambda_3(\tau)$
6A	3	-3	$-9\Lambda_2(\tau) - 2\Lambda_3(\tau) + 3\Lambda_4(\tau) + 3\Lambda_6(\tau) - \Lambda_{12}(\tau)$
3B	0	0	$8\Lambda_3(\tau) - 2\Lambda_9(\tau) + 2\,\eta^6(\tau)/\eta^2(3\tau)$
6B	0	0	$-2\eta(\tau)^5\eta(3\tau)/\eta(2\tau)\eta(6\tau)$
4B	0	0	$-2\eta(2\tau)^8/\eta(4\tau)^4$
4C	4	0	$-8\Lambda_4(\tau)/3$
5A	2	2	$-2\Lambda_5(\tau)$
10A	2	-2	$\sum_{d\mid 20} c_{10A}(d)\Lambda_d(\tau) + \frac{20}{3}\eta(2\tau)^2\eta(10\tau)^2$
12A	0	0	$-2\eta(\tau)\eta(2\tau)^5\eta(3\tau)/\eta(4\tau)^2\eta(6\tau)$
6C	1	1	$2(\Lambda_2(\tau) + \Lambda_3(\tau) - \Lambda_6(\tau))$
6D	1	-1	$-5\Lambda_2(\tau) - 2\Lambda_3(\tau) + \frac{5}{3}\Lambda_4(\tau) + 3\Lambda_6(\tau) - \Lambda_{12}(\tau)$
8AB	0	0	$-2\eta(2\tau)^4\eta(4\tau)^2/\eta(8\tau)^2$
8CD	2	0	$-2\Lambda_2(\tau) + \frac{5}{3}\Lambda_4(\tau) - \Lambda_8(\tau)$
20AB	0	0	$-2\eta(2\tau)^7\eta(5\tau)/\eta(\tau)\eta(4\tau)^2\eta(10\tau)$
11AB	1	1	$-\frac{2}{5}\Lambda_{11}(\tau) - \frac{33}{5}\eta(\tau)^2\eta(11\tau)^2$
22AB	1	-1	$\sum_{d\mid 44} c_g(d)\Lambda_d(\tau) - \frac{11}{5}\sum_{d\mid 4} c'_g(d)\eta(d\tau)^2\eta(11d\tau)^2 + \frac{22}{3}f_{44}(\tau)$

The function f_{44} is the unique new cusp form of weight 2 for $\Gamma_0(44)$, normalized so that $f_{44}(\tau) =$

$q+O(q^3)$ as $\Im(\tau)\to\infty$. The coefficients $c_g(d)$ and $c'_g(d)$ for $g\in 10A\cup 22A\cup 22B$ are given by

$$c_{10A}(2)=-5,\ c_{10A}(4)=-\frac{5}{3},\ c_{10A}(5)=-\frac{2}{3},\ c_{10A}(10)=1,\ c_{10A}(20)=-\frac{1}{3}, \tag{B.30}$$

$$c_{22AB}(2)=-\frac{11}{5},\ c_{22AB}(4)=\frac{11}{5},\ c_{22AB}(11)=-\frac{2}{15},\ c_{22AB}(22)=\frac{1}{5},\ c_{22AB}(44)=-\frac{1}{15}, \tag{B.31}$$

$$c'_{22AB}(1)=1,\ c'_{22AB}(2)=4,\ c'_{22AB}(4)=8. \tag{B.32}$$

B.3.3 $\ell=4$, $X=A_3^8$

We have $m^X=4$, so the umbral McKay-Thompson series $H_g^{(4)}=(H_{g,r}^{(4)})$ associated to $g\in G^{(4)}$ is an 8-vector-valued function, with components indexed by $r\in\mathbb{Z}/8\mathbb{Z}$.

Define $H_g^{(4)}=(H_{g,r}^{(4)})$ for $g\in G^{(4)}$, $g\notin 4C$, by requiring that

$$\psi_g^{(4)}(\tau,z)=-\chi_g^{(4)}\mu_{4,0}^0(\tau,z)-\bar{\chi}_g^{(4)}\mu_{4,0}^1(\tau,z)+\sum_{r \bmod 8}H_{g,r}^{(4)}(\tau)\theta_{4,r}(\tau,z), \tag{B.33}$$

where $\chi_g^{(4)}:=\chi_g^{X_A}$ and $\bar{\chi}_g^{(4)}:=\bar{\chi}_g^{X_A}$ (cf. Table 17), and the $\psi_g^{(4)}$ are meromorphic Jacobi forms of weight 1 and index 4 given explicitly in Table 38.

Table 38: Character Values and Meromorphic Jacobi Forms for $\ell=4$, $X=A_3^8$

$[g]$	$\chi_g^{(4)}$	$\bar{\chi}_g^{(4)}$	$\psi_g^{(4)}(\tau,z)$
1A	8	8	$2i\theta_1(\tau,2z)^3\theta_1(\tau,z)^{-4}\eta(\tau)^3$
2A	-8	8	$2i\theta_1(\tau,2z)^3\theta_2(\tau,z)^{-4}\eta(\tau)^3$
2B	0	0	$-2i\theta_1(\tau,2z)^3\theta_1(\tau,z)^{-2}\theta_2(\tau,z)^{-2}\eta(\tau)^3$
4A	0	0	$-2i\theta_1(\tau,2z)\theta_2(\tau,2z)^2\theta_2(2\tau,2z)^{-2}\eta(2\tau)^2\eta(\tau)^{-1}$
4B	0	0	$-2i\theta_1(2\tau,2z)\theta_3(2\tau,2z)^2\theta_4(2\tau,2z)\eta(2\tau)^2\eta(\tau)^{-2}\eta(4\tau)^{-2}$
2C	0	4	$2i\theta_1(\tau,2z)\theta_2(\tau,2z)^2\theta_1(\tau,z)^{-2}\theta_2(\tau,z)^{-2}\eta(\tau)^3$
3A	2	2	$2i\theta_1(3\tau,6z)\theta_1(\tau,z)^{-1}\theta_1(3\tau,3z)^{-1}\eta(\tau)^3$
6A	-2	2	$-2i\theta_1(3\tau,6z)\theta_2(\tau,z)^{-1}\theta_2(3\tau,3z)^{-1}\eta(\tau)^3$
6BC	0	0	cf. (B.34)
8A	0	0	$-2i\theta_1(\tau,2z)\theta_2(2\tau,4z)\theta_2(4\tau,4z)^{-1}\eta(\tau)\eta(4\tau)\eta(2\tau)^{-1}$
4C	0	2	$2i\theta_1(\tau,2z)\theta_2(2\tau,4z)\theta_1(2\tau,2z)^{-2}\eta(2\tau)^7\eta(\tau)^{-3}\eta(4\tau)^{-2}$
7AB	1	1	cf. (B.34)
14AB	-1	1	cf. (B.34)

$$\begin{aligned}
\psi^{(4)}_{6BC} &:= \left(\theta_1(\tau, z+\tfrac{1}{3})\theta_1(\tau, z+\tfrac{1}{6}) - \theta_1(\tau, z-\tfrac{1}{3})\theta_1(\tau, z-\tfrac{1}{6})\right) \frac{-i\theta_1(3\tau, 6z)}{\theta_1(3\tau, 3z)\theta_2(3\tau, 3z)}\eta(3\tau) \\
\psi^{(4)}_{7AB} &:= \left(\prod_{j=1}^{3}\theta_1(\tau, 2z+\tfrac{j^2}{7})\theta_1(\tau, z-\tfrac{j^2}{7}) + \prod_{j=1}^{3}\theta_1(\tau, 2z-\tfrac{j^2}{7})\theta_1(\tau, z+\tfrac{j^2}{7})\right) \frac{-i}{\theta_1(7\tau, 7z)}\frac{\eta(7\tau)}{\eta(\tau)^4} \\
\psi^{(4)}_{14AB} &:= \left(\prod_{j=1}^{3}\theta_1(\tau, 2z+\tfrac{j^2}{7})\theta_2(\tau, z-\tfrac{j^2}{7}) + \prod_{j=1}^{3}\theta_1(\tau, 2z-\tfrac{j^2}{7})\theta_2(\tau, z+\tfrac{j^2}{7})\right) \frac{i}{\theta_2(7\tau, 7z)}\frac{\eta(7\tau)}{\eta(\tau)^4}
\end{aligned} \tag{B.34}$$

For use later on, note that $\psi^{(4)}_{1A} = -2\Psi_{1,1}\varphi^{(4)}_1$, where

$$\varphi^{(4)}_1(\tau, z) := \frac{\theta_1(\tau, 2z)^2}{\theta_1(\tau, z)^2}. \tag{B.35}$$

B.3.4 $\ell = 5$, $X = A_4^6$

We have $m^X = 5$, so the umbral McKay-Thompson series $H^{(5)}_g = (H^{(5)}_{g,r})$ associated to $g \in G^{(5)}$ is a 10-vector-valued function, with components indexed by $r \in \mathbb{Z}/10\mathbb{Z}$.

Define $H^{(5)}_g = (H^{(5)}_{g,r})$ for $g \in G^{(5)}$, $g \notin 5A \cup 10A$, by requiring that

$$\psi^{(5)}_g(\tau, z) = -\chi^{(5)}_g \mu^0_{5,0}(\tau, z) - \bar{\chi}^{(5)}_g \mu^1_{5,0}(\tau, z) + \sum_{r \bmod 10} H^{(5)}_{g,r}(\tau)\theta_{5,r}(\tau, z), \tag{B.36}$$

where $\chi^{(5)}_g := \chi^{X_A}_g$ and $\bar{\chi}^{(5)}_g := \bar{\chi}^{X_A}_g$ (cf. Table 18), and the $\psi^{(5)}_g$ are meromorphic Jacobi forms of weight 1 and index 5 given explicitly in Table 39.

Table 39: Character Values and Meromorphic Jacobi Forms for $\ell = 5$, $X = A_4^6$

$[g]$	$\chi^{(5)}_g$	$\bar{\chi}^{(5)}_g$	$\psi^{(5)}_g(\tau, z)$
1A	6	6	$2i\theta_1(\tau, 2z)\theta_1(\tau, 3z)\theta_1(\tau, z)^{-3}\eta(\tau)^3$
2A	-6	6	$-2i\theta_1(\tau, 2z)\theta_2(\tau, 3z)\theta_2(\tau, z)^{-3}\eta(\tau)^3$
2B	-2	2	$-2i\theta_1(\tau, 2z)\theta_1(\tau, 3z)\theta_1(\tau, z)^{-1}\theta_2(\tau, z)^{-2}\eta(\tau)^3$
2C	2	2	$2i\theta_1(\tau, 2z)\theta_2(\tau, 3z)\theta_1(\tau, z)^{-2}\theta_2(\tau, z)^{-1}\eta(\tau)^3$
3A	0	0	$-2i\theta_1(\tau, 2z)\theta_1(\tau, 3z)\theta_1(3\tau, 3z)^{-1}\eta(3\tau)$
6A	0	0	$-2i\theta_1(\tau, 2z)\theta_2(\tau, 3z)\theta_2(3\tau, 3z)^{-1}\eta(3\tau)$
4AB	0	0	cf. (B.37)
4CD	0	2	cf. (B.37)
12AB	0	0	cf. (B.37)

$$\begin{aligned}\psi^{(5)}_{4AB}(\tau,z)&:=-i\theta_2(\tau,2z)\frac{\theta_1(\tau,z+\frac14)\theta_1(\tau,3z+\frac14)-\theta_1(\tau,z-\frac14)\theta_1(\tau,3z-\frac14)}{\theta_2(2\tau,2z)^2}\frac{\eta(2\tau)^2}{\eta(\tau)}\\ \psi^{(5)}_{4CD}(\tau,z)&:=-i\theta_2(\tau,2z)\frac{\theta_1(\tau,z+\frac14)\theta_1(\tau,3z-\frac14)+\theta_1(\tau,z-\frac14)\theta_1(\tau,3z+\frac14)}{\theta_1(2\tau,2z)\theta_2(2\tau,2z)}\frac{\eta(2\tau)^2}{\eta(\tau)}\\ \psi^{(5)}_{12AB}(\tau,z)&:=i\frac{\theta_2(\tau,2z)}{\theta_2(6\tau,6z)}\left(\theta_1(\tau,z+\tfrac{1}{12})\theta_1(\tau,z+\tfrac14)\theta_1(\tau,z+\tfrac{5}{12})\theta_1(\tau,3z-\tfrac14)\right.\\ &\qquad\left.-\theta_1(\tau,z-\tfrac{1}{12})\theta_1(\tau,z-\tfrac14)\theta_1(\tau,z-\tfrac{5}{12})\theta_1(\tau,3z+\tfrac14)\right)\frac{\eta(6\tau)}{\eta(\tau)^3}\end{aligned}\tag{B.37}$$

For $g\in 5A$ use the formulas of §B.3.20 to define

$$H^{(5)}_{5A,r}(\tau):=H^{(25)}_{1A,r}(\tau/5)-H^{(25)}_{1A,10-r}(\tau/5)+H^{(25)}_{1A,10+r}(\tau/5)-H^{(25)}_{1A,20-r}(\tau/5)+H^{(25)}_{1A,20+r}(\tau/5).\tag{B.38}$$

For $g\in 10A$ set $H^{(5)}_{10A,r}(\tau):=-(-1)^rH^{(5)}_{5A,r}(\tau)$.

For use later on we note that $\psi^{(5)}_{1A}=-2\Psi_{1,1}\varphi^{(5)}_1$, where

$$\varphi^{(5)}_1(\tau,z):=\frac{\theta_1(\tau,3z)}{\theta_1(\tau,z)}.\tag{B.39}$$

B.3.5 $\ell=6$, $X=A_5^4D_4$

We have $m^X=6$, so the umbral McKay-Thompson series $H^{(6)}_g=(H^{(6)}_{g,r})$ associated to $g\in G^{(6)}$ is a 12-vector-valued function with components indexed by $r\in\mathbb{Z}/12\mathbb{Z}$. We have $H^{(6)}_{g,r}=-H^{(6)}_{g,-r}$, so it suffices to specify the $H^{(6)}_{g,r}$ for $r\in\{1,2,3,4,5\}$.

To define $H^{(6)}_g=(H^{(6)}_{g,r})$ for $g=e$, first define $h(\tau)=(h_r(\tau))$ by requiring that

$$-2\Psi_{1,1}(\tau,z)\varphi^{(6)}_1(\tau,z)=-24\mu_{6,0}(\tau,z)+\sum_{r\bmod 12}h_r(\tau)\theta_{6,r}(\tau,z),\tag{B.40}$$

where

$$\varphi^{(6)}_1(\tau,z):=\varphi^{(2)}_1(\tau,z)\varphi^{(5)}_1(\tau,z)-\varphi^{(3)}_1(\tau,z)\varphi^{(4)}_1(\tau,z).\tag{B.41}$$

(Cf. (B.23), (B.27), (B.35), (B.39).) Now define the $H^{(6)}_{1A,r}$ by setting

$$\begin{aligned}H^{(6)}_{1A,1}(\tau)&:=\frac{1}{24}\left(5h_1(\tau)+h_5(\tau)\right),\\ H^{(6)}_{1A,2}(\tau)&:=\frac16h_2(\tau),\\ H^{(6)}_{1A,3}(\tau)&:=\frac14h_3(\tau),\\ H^{(6)}_{1A,4}(\tau)&:=\frac16h_4(\tau),\\ H^{(6)}_{1A,5}(\tau)&:=\frac{1}{24}\left(h_1(\tau)+5h_5(\tau)\right).\end{aligned}\tag{B.42}$$

Define $H^{(6)}_{2A,r}$ by requiring

$$H^{(6)}_{2A,r}(\tau) := -(-1)^r H^{(6)}_{1A,r}(\tau). \tag{B.43}$$

For the remaining g, recall (B.13). The $H^{(6)}_{g,r}$ for $g \notin 1A \cup 2A$ are defined as follows for $r = 2$ and $r = 4$, noting that $H^{(3)}_{g,4} = H^{(3)}_{g,-2} = -H^{(3)}_{g,2}$.

$$\begin{aligned}
H^{(6)}_{2B,r}(\tau) &:= [-\tfrac{r^2}{24}]H^{(3)}_{4C,r}(\tau/2)\\
H^{(6)}_{4A,r}(\tau) &:= [-\tfrac{r^2}{24}]H^{(3)}_{4B,r}(\tau/2)\\
H^{(6)}_{3A,r}(\tau) &:= [-\tfrac{r^2}{24}]H^{(3)}_{6C,r}(\tau/2)\\
H^{(6)}_{6A,r}(\tau) &:= [-\tfrac{r^2}{24}]H^{(3)}_{6D,r}(\tau/2)\\
H^{(6)}_{8AB,r}(\tau) &:= [-\tfrac{r^2}{24}]H^{(3)}_{8CD,r}(\tau/2)
\end{aligned} \tag{B.44}$$

For the $H^{(6)}_{g,3}$ we define

$$\begin{aligned}
H^{(6)}_{2B,3}(\tau), H^{(6)}_{4A,3}(\tau) &:= -[-\tfrac{9}{24}]H^{(2)}_{6A,1}(\tau/3),\\
H^{(6)}_{3A,3}(\tau), H^{(6)}_{6A,3}(\tau) &:= 0,\\
H^{(6)}_{8AB,3}(\tau) &:= -[-\tfrac{9}{24}]H^{(2)}_{12A,1}(\tau/3).
\end{aligned} \tag{B.45}$$

Noting that $H^{(2)}_{g,5} = H^{(2)}_{g,1}$ and $H^{(3)}_{g,5} = -H^{(3)}_{g,1}$, the $H^{(6)}_{g,1}$ and $H^{(6)}_{g,5}$ are defined for $o(g) \neq 0 \mod 3$ by setting

$$\begin{aligned}
H^{(6)}_{2B,r}(\tau) &:= [-\tfrac{1}{24}]\frac{1}{2}\left(H^{(2)}_{6A,r}(\tau/3) + H^{(3)}_{4C,r}(\tau/2)\right)\\
H^{(6)}_{4A,r}(\tau) &:= [-\tfrac{1}{24}]\frac{1}{2}\left(H^{(2)}_{6A,r}(\tau/3) + H^{(3)}_{4B,r}(\tau/2)\right)\\
H^{(6)}_{8AB,r}(\tau) &:= [-\tfrac{1}{24}]\frac{1}{2}\left(H^{(2)}_{12A,r}(\tau/3) + H^{(3)}_{8CD,r}(\tau/2)\right)
\end{aligned} \tag{B.46}$$

It remains to specify the $H^{(6)}_{g,r}$ when $g \in 3A \cup 6A$ and r is 1 or 5. These cases are determined by using the formulas of §B.3.17 to set

$$\begin{aligned}
H^{(6)}_{3A,1}(\tau), H^{(6)}_{6A,1}(\tau) &:= H^{(18)}_{1A,1}(3\tau) - H^{(18)}_{1A,11}(3\tau) + H^{(18)}_{1A,13}(3\tau),\\
H^{(6)}_{3A,5}(\tau), H^{(6)}_{6A,5}(\tau) &:= H^{(18)}_{1A,5}(3\tau) - H^{(18)}_{1A,7}(3\tau) + H^{(18)}_{1A,17}(3\tau).
\end{aligned} \tag{B.47}$$

B.3.6 $\ell = 6+3$, $X = D_4^6$

We have $m^X = 6$, so the umbral McKay-Thompson series $H^{(6+3)}_g = (H^{(6+3)}_{g,r})$ associated to $g \in G^{(6+3)}$ is a 12-vector-valued function with components indexed by $r \in \mathbb{Z}/12\mathbb{Z}$. In addition to the identity $H^{(6+3)}_{g,r} = -H^{(6+3)}_{g,-r}$, we have $H^{(6+3)}_{g,r} = 0$ for $r = 0 \mod 2$. Thus it suffices to specify the $H^{(6+3)}_{g,r}$ for $r \in \{1, 3, 5\}$.

Recall (B.13). For $r=1$, define

$$\begin{aligned}
H^{(6+3)}_{1A,1}(\tau), H^{(6+3)}_{3A,1}(\tau) &:= H^{(6)}_{1A,1}(\tau) + H^{(6)}_{1A,5}(\tau),\\
H^{(6+3)}_{2A,1}(\tau), H^{(6+3)}_{6A,1}(\tau) &:= H^{(6)}_{2B,1}(\tau) + H^{(6)}_{2B,5}(\tau),\\
H^{(6+3)}_{3B,1}(\tau) &:= H^{(6)}_{3A,1}(\tau) + H^{(6)}_{3A,5}(\tau),\\
H^{(6+3)}_{3C,1}(\tau) &:= -2\frac{\eta(\tau)^2}{\eta(3\tau)},\\
H^{(6+3)}_{4A,1}(\tau), H^{(6+3)}_{12A,1}(\tau) &:= H^{(6)}_{8AB,1}(\tau) + H^{(6)}_{8AB,5}(\tau),\\
H^{(6+3)}_{5A,1}(\tau), H^{(6+3)}_{15A,1}(\tau) &:= [-\tfrac{1}{24}]H^{(2)}_{15AB,1}(\tau/3),\\
H^{(6+3)}_{2C,1}(\tau) &:= H^{(6)}_{4A,1}(\tau) - H^{(6)}_{4A,5}(\tau),\\
H^{(6+3)}_{4B,1}(\tau) &:= H^{(6)}_{8AB,1}(\tau) - H^{(6)}_{8AB,5}(\tau),\\
H^{(6+3)}_{6B,1}(\tau) &:= H^{(6)}_{6A,1}(\tau) - H^{(6)}_{6A,5}(\tau),\\
H^{(6+3)}_{6C,1}(\tau) &:= -2\frac{\eta(2\tau)\,\eta(3\tau)}{\eta(6\tau)}.
\end{aligned} \tag{B.48}$$

Then define $H^{(6+3)}_{2B,1}$ by setting

$$H^{(6+3)}_{2B,1}(\tau) := 2H^{(6+3)}_{4B,1}(\tau) + 2\frac{\eta(\tau)^3}{\eta(2\tau)^2}. \tag{B.49}$$

For $r=3$ set

$$\begin{aligned}
H^{(6+3)}_{1A,3}(\tau) &:= 2H^{(6)}_{1A,3}(\tau),\\
H^{(6+3)}_{3A,3}(\tau) &:= -H^{(6)}_{1A,3}(\tau),\\
H^{(6+3)}_{2A,3}(\tau) &:= 2H^{(6)}_{2B,3}(\tau),\\
H^{(6+3)}_{6A,3}(\tau) &:= -H^{(6)}_{2B,3}(\tau),\\
H^{(6+3)}_{4A,3}(\tau) &:= 2H^{(6)}_{8AB,3}(\tau),\\
H^{(6+3)}_{12A,3}(\tau) &:= -H^{(6)}_{8AB,3}(\tau),\\
H^{(6+3)}_{5A,3}(\tau) &:= -2[-\tfrac{9}{24}]H^{(2)}_{15AB,1}(\tau),\\
H^{(6+3)}_{15A,3}(\tau) &:= [-\tfrac{9}{24}]H^{(2)}_{15AB,1}(\tau),
\end{aligned} \tag{B.50}$$

and

$$H^{(6+3)}_{3B,3}(\tau), H^{(6+3)}_{3C,3}(\tau), H^{(6+3)}_{2B,3}(\tau), H^{(6+3)}_{2C,3}(\tau), H^{(6+3)}_{4B,3}(\tau), H^{(6+3)}_{6B,3}(\tau), H^{(6+3)}_{6C,3}(\tau) := 0. \tag{B.51}$$

For $r=5$ define $H^{(6+3)}_{g,5}(\tau) := H^{(6+3)}_{g,1}(\tau)$ for $[g] \in \{1A, 3A, 2A, 6A, 3B, 3C, 4A, 12A, 5A, 15AB\}$, and set $H^{(6+3)}_{g,5}(\tau) := -H^{(6+3)}_{g,1}(\tau)$ for the remaining cases, $[g] \in \{2B, 2C, 4B, 6B, 6C\}$.

B.3.7 $\ell = 7$, $X = A_6^4$

We have $m^X = 7$, so the umbral McKay-Thompson series $H_g^{(7)} = (H_{g,r}^{(7)})$ associated to $g \in G^{(7)} = G^X \simeq SL_2(3)$ is a 14-vector-valued function, with components indexed by $r \in \mathbb{Z}/14\mathbb{Z}$.

Define $H_g^{(7)} = (H_{g,r}^{(7)})$ for $g \in G^{(7)}$ by requiring that

$$\psi_g^{(7)}(\tau, z) = -\chi_g^{(7)}\mu_{7,0}^0(\tau, z) - \bar{\chi}_g^{(7)}\mu_{7,0}^1(\tau, z) + \sum_{r \bmod 14} H_{g,r}^{(7)}(\tau)\theta_{7,r}(\tau, z), \tag{B.52}$$

where $\chi_g^{(7)} := \chi_g^{X_A}$ and $\bar{\chi}_g^{(7)} := \bar{\chi}_g^{X_A}$ (cf. Table 21), and the $\psi_g^{(7)}$ are meromorphic Jacobi forms of weight 1 and index 7 given explicitly in Table 40.

Table 40: Character Values and Meromorphic Jacobi Forms for $\ell = 7$, $X = A_6^4$

$[g]$	$\chi_g^{(7)}$	$\bar{\chi}_g^{(7)}$	$\psi_g^{(7)}(\tau, z)$
1A	4	4	$2i\theta_1(\tau, 4z)\theta_1(\tau, z)^{-2}\eta(\tau)^3$
2A	-4	4	$-2i\theta_1(\tau, 4z)\theta_2(\tau, z)^{-2}\eta(\tau)^3$
4A	0	0	$-2i\theta_1(\tau, 4z)\theta_2(2\tau, 2z)^{-1}\eta(2\tau)\eta(\tau)$
3A	1	1	cf. (B.53)
6A	-1	1	cf. (B.53)

$$\begin{aligned}
\psi_{3A}^{(7)}(\tau, z) &:= -i\frac{\theta_1(\tau, 4z + \frac{1}{3})\theta_1(\tau, z - \frac{1}{3}) + \theta_1(\tau, 4z - \frac{1}{3})\theta_1(\tau, z + \frac{1}{3})}{\theta_1(3\tau, 3z)}\eta(3\tau) \\
\psi_{6A}^{(7)}(\tau, z) &:= -i\frac{\theta_1(\tau, 4z + \frac{1}{3})\theta_1(\tau, z - \frac{1}{6}) - \theta_1(\tau, 4z - \frac{1}{3})\theta_1(\tau, z + \frac{1}{6})}{\theta_2(3\tau, 3z)}\eta(3\tau)
\end{aligned} \tag{B.53}$$

For use later on we note that $\psi_{1A}^{(7)} = -2\Psi_{1,1}\varphi_1^{(7)}$, where

$$\varphi_1^{(7)}(\tau, z) := \frac{\theta_1(\tau, 4z)}{\theta_1(\tau, 2z)}. \tag{B.54}$$

B.3.8 $\ell = 8$, $X = A_7^2 D_5^2$

We have $m^X = 8$, so the umbral McKay-Thompson series $H_g^{(8)} = (H_{g,r}^{(8)})$ associated to $g \in G^{(8)}$ is a 16-vector-valued function with components indexed by $r \in \mathbb{Z}/16\mathbb{Z}$. We have $H_{g,r}^{(8)} = -H_{g,-r}^{(8)}$, so it suffices to specify the $H_{g,r}^{(8)}$ for $r \in \{1, 2, 3, 4, 5, 6, 7\}$.

To define $H_g^{(8)} = (H_{g,r}^{(8)})$ for $g = e$, first define $h(\tau) = (h_r(\tau))$ by requiring that

$$-2\Psi_{1,1}(\tau, z)\left(\varphi_1^{(8)}(\tau, z) + \frac{1}{2}\varphi_2^{(8)}(\tau, z)\right) = -24\mu_{8,0}(\tau, z) + \sum_{r \bmod 16} h_r(\tau)\theta_{8,r}(\tau, z), \tag{B.55}$$

where

$$\begin{aligned}
\varphi_1^{(8)}(\tau, z) &:= \varphi_1^{(3)}(\tau, z)\varphi_1^{(6)}(\tau, z) - 5\varphi_1^{(4)}(\tau, z)\varphi_1^{(5)}(\tau, z), \\
\varphi_2^{(8)}(\tau, z) &:= \varphi_1^{(4)}(\tau, z)\varphi_1^{(5)}(\tau, z) - \varphi_1^{(8)}(\tau, z).
\end{aligned} \tag{B.56}$$

(Cf. (B.27), (B.35), (B.39), (B.41).) Now define the $H^{(8)}_{1A,r}$ by setting

$$H^{(8)}_{1A,r}(\tau) := \frac{1}{6}h_r(\tau), \tag{B.57}$$

for $r \in \{1,3,4,5,7\}$, and

$$H^{(8)}_{1A,2}(\tau), H^{(8)}_{1A,6}(\tau) := \frac{1}{12}\left(h_2(\tau) + h_6(\tau)\right). \tag{B.58}$$

Define $H^{(8)}_{2A,r}$ for $1 \leq r \leq 7$ by requiring

$$H^{(8)}_{2A,r}(\tau) := -(-1)^r H^{(8)}_{1A,r}(\tau). \tag{B.59}$$

For the remaining g, recall (B.13). The $H^{(8)}_{g,r}$ for $g \in 2B \cup 2C \cup 4A$ are defined as follows for $r \in \{1,3,5,7\}$, noting that $H^{(4)}_{g,7} = H^{(4)}_{g,-1} = -H^{(4)}_{g,1}$, &c.

$$\begin{aligned} H^{(8)}_{2BC,r}(\tau) &:= [-\tfrac{r^2}{32}]H^{(4)}_{4C,r}(\tau/2) \\ H^{(8)}_{4A,r}(\tau) &:= [-\tfrac{r^2}{32}]H^{(4)}_{4B,r}(\tau/2) \end{aligned} \tag{B.60}$$

The $H^{(8)}_{2BC,r}$ and $H^{(8)}_{4A,r}$ vanish for $r = 0 \mod 2$.

B.3.9 $\ell = 9$, $X = A_8^3$

We have $m^X = 9$, so for $g \in G^{(9)}$ the associated umbral McKay-Thompson series $H^{(9)}_g = (H^{(9)}_{g,r})$ is a 18-vector-valued function, with components indexed by $r \in \mathbb{Z}/18\mathbb{Z}$, satisfying $H^{(9)}_{g,r} = -H^{(9)}_{g,-r}$, and in particular, $H^{(9)}_{g,r} = 0$ for $r = 0 \mod 9$. So it suffices to specify the $H^{(9)}_{g,r}$ for $r \in \{1,2,3,4,5,6,7,8\}$.

Define $H^{(9)}_g = (H^{(9)}_{g,r})$ for $g = e$ by requiring that

$$-\Psi_{1,1}(\tau,z)\varphi^{(9)}_1(\tau,z) = -3\mu_{9,0}(\tau,z) + \sum_{r \mod 18} H^{(9)}_{e,r}(\tau)\theta_{9,r}(\tau,z), \tag{B.61}$$

where

$$\varphi^{(9)}_1(\tau,z) := \varphi^{(3)}_1(\tau,z)\varphi^{(7)}_1(\tau,z) - \varphi^{(5)}_1(\tau,z)^2. \tag{B.62}$$

(Cf. (B.27), (B.39), (B.54).)

Recall (B.13). The $H^{(9)}_{2B,r}$ are defined for $r \in \{1,2,4,5,7,8\}$ by setting

$$H^{(9)}_{2B,r}(\tau) := [-\tfrac{r^2}{36}]H^{(3)}_{6C,r}(\tau/3), \tag{B.63}$$

where we note that $H^{(3)}_{g,4} = H^{(3)}_{g,-2} = -H^{(3)}_{g,2}$, &c. We determine $H^{(9)}_{2B,3}$ and $H^{(9)}_{2B,6}$ by using §B.3.17 to set

$$H^{(9)}_{2B,r}(\tau) := H^{(18)}_{1A,r}(2\tau) - H^{(18)}_{1A,18-r}(2\tau) \tag{B.64}$$

for $r \in \{3,6\}$.

The $H^{(9)}_{3A,r}$ are defined by the explicit formulas

$$\begin{aligned}
H^{(9)}_{3A,1}(\tau) &:= [-\tfrac{1}{36}]f^{(9)}_1(\tau/3),\\
H^{(9)}_{3A,2}(\tau) &:= [-\tfrac{4}{36}]f^{(9)}_2(\tau/3),\\
H^{(9)}_{3A,3}(\tau) &:= -\theta_{3,3}(\tau,0),\\
H^{(9)}_{3A,4}(\tau) &:= -[-\tfrac{16}{36}]f^{(9)}_2(\tau/3),\\
H^{(9)}_{3A,5}(\tau) &:= -[-\tfrac{25}{36}]f^{(9)}_1(\tau/3),\\
H^{(9)}_{3A,6}(\tau) &:= \theta_{3,0}(\tau,0),\\
H^{(9)}_{3A,7}(\tau) &:= [-\tfrac{13}{36}]f^{(9)}_1(\tau/3),\\
H^{(9)}_{3A,8}(\tau) &:= [-\tfrac{28}{36}]f^{(9)}_2(\tau/3),
\end{aligned} \tag{B.65}$$

where

$$\begin{aligned}
f^{(9)}_1(\tau) &:= -2\frac{\eta(\tau)\eta(12\tau)\eta(18\tau)^2}{\eta(6\tau)\eta(9\tau)\eta(36\tau)},\\
f^{(9)}_2(\tau) &:= \frac{\eta(2\tau)^6\eta(12\tau)\eta(18\tau)^2}{\eta(\tau)\eta(4\tau)^4\eta(6\tau)\eta(9\tau)\eta(36\tau)} - \frac{\eta(\tau)\eta(2\tau)\eta(3\tau)^2}{\eta(4\tau)^2\eta(9\tau)}.
\end{aligned} \tag{B.66}$$

Finally, the $H^{(9)}_{g,r}$ are determined for $g \in 2A \cup 2C \cup 6A$ by setting

$$\begin{aligned}
H^{(9)}_{2A,r}(\tau) &:= (-1)^{r+1}H^{(9)}_{1A,r}(\tau),\\
H^{(9)}_{2C,r}(\tau) &:= (-1)^{r+1}H^{(9)}_{2B,r}(\tau),\\
H^{(9)}_{6A,r}(\tau) &:= (-1)^{r+1}H^{(9)}_{3A,r}(\tau).
\end{aligned} \tag{B.67}$$

B.3.10 $\ell = 10$, $X = A_9^2D_6$

We have $m^X = 10$, so the umbral McKay-Thompson series $H^{(10)}_g = (H^{(10)}_{g,r})$ associated to $g \in G^{(10)}$ is a 20-vector-valued function with components indexed by $r \in \mathbb{Z}/20\mathbb{Z}$. We have $H^{(10)}_{g,r} = -H^{(10)}_{g,-r}$, so it suffices to specify the $H^{(10)}_{g,r}$ for $1 \leq r \leq 9$.

To define $H^{(10)}_g = (H^{(10)}_{g,r})$ for $g = e$, first define $h(\tau) = (h_r(\tau))$ by requiring that

$$-6\Psi_{1,1}(\tau,z)\varphi^{(10)}_1(\tau,z) = -24\mu_{10,0}(\tau,z) + \sum_{r \mod 20} h_r(\tau)\theta_{10,r}(\tau,z), \tag{B.68}$$

where

$$\varphi^{(10)}_1(\tau,z) := 5\varphi^{(4)}_1(\tau,z)\varphi^{(7)}_1(\tau,z) - \varphi^{(5)}_1(\tau,z)\varphi^{(6)}_1(\tau,z). \tag{B.69}$$

(Cf. (B.35), (B.39), (B.41), (B.54).) Now define the $H^{(10)}_{1A,r}$ for r odd by setting

$$\begin{aligned} H^{(10)}_{1A,1}(\tau) &:= \frac{1}{24}\left(3h_1(\tau)+h_9(\tau)\right),\\ H^{(10)}_{1A,3}(\tau) &:= \frac{1}{24}\left(3h_3(\tau)+h_7(\tau)\right),\\ H^{(10)}_{1A,5}(\tau) &:= \frac{1}{6}h_5(\tau),\\ H^{(10)}_{1A,3}(\tau) &:= \frac{1}{24}\left(h_3(\tau)+3h_7(\tau)\right),\\ H^{(10)}_{1A,9}(\tau) &:= \frac{1}{24}\left(h_1(\tau)+3h_9(\tau)\right). \end{aligned} \tag{B.70}$$

For $r = 0 \mod 2$ set

$$H^{(10)}_{1A,r}(\tau) := \frac{1}{12}h_r(\tau), \tag{B.71}$$

and define $H^{(10)}_{2A,r}$ for $1 \leq r \leq 9$ by requiring

$$H^{(10)}_{2A,r}(\tau) := -(-1)^r H^{(10)}_{1A,r}(\tau). \tag{B.72}$$

It remains to specify $H^{(10)}_{g,r}$ for $g \in 4A \cup 4B$. For $r = 0 \mod 2$ set

$$H^{(10)}_{4AB,r}(\tau) := 0. \tag{B.73}$$

For r odd, recall (B.13), and define

$$H^{(10)}_{4A,r}(\tau) := [-\tfrac{r^2}{40}]\frac{1}{2}\left(H^{(2)}_{10A,r}(\tau/5) + H^{(5)}_{4CD,r}(\tau/2)\right). \tag{B.74}$$

B.3.11 $\ell = 10+5$, $X = D_6^4$

We have $m^X = 10$, so the umbral McKay-Thompson series $H^{(10+5)}_g = (H^{(10+5)}_{g,r})$ associated to $g \in G^{(10+5)}$ is a 20-vector-valued function with components indexed by $r \in \mathbb{Z}/20\mathbb{Z}$. We have $H^{(10+5)}_{g,r} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H^{(10+5)}_{g,r}$ for r odd. Observing that $H^{(10+5)}_{g,r} = -H^{(10+5)}_{g,-r}$ we may determine $H^{(10+5)}_g$ by requiring that

$$\psi^{(5/2)}_g(\tau,z) = -2\chi^{(5/2)}_g i\mu_{5/2,0}(\tau,z) + \sum_{\substack{r\in\mathbb{Z}+1/2\\ r \mod 5}} e(-r/2)H^{(10+5)}_{g,2r}(\tau)\theta_{5/2,r}(\tau,z), \tag{B.75}$$

where $\chi^{(5/2)}_g := \bar{\chi}^{X_D}_g$ as in Table 25, and the $\psi^{(5/2)}_g$ are the meromorphic Jacobi forms of weight 1 and index 5/2 defined as follows.

Table 41: Character Values and Meromorphic Jacobi Forms for $\ell = 10 + 5$, $X = D_6^4$

$[g]$	$\bar{\chi}_g^{(5/2)}$	$\psi_g^{(5/2)}(\tau, z)$
1A	4	$2i\theta_1(\tau, 2z)^2\theta_1(\tau, z)^{-3}\eta(\tau)^3$
2A	0	$-2i\theta_1(\tau, 2z)^2\theta_1(\tau, z)^{-1}\theta_2(\tau, z)^{-2}\eta(\tau)^3$
3A	1	$2i\theta_1(3\tau, 6z)\theta_1(\tau, 2z)^{-1}\theta_1(3\tau, 3z)^{-1}\eta(\tau)^3$
2B	2	$2i\theta_1(\tau, 2z)\theta_2(\tau, 2z)\theta_1(\tau, z)^{-2}\theta_2(\tau, z)^{-1}\eta(\tau)^3$
4A	0	$-2i\theta_1(\tau, 2z)\theta_2(\tau, 2z)\theta_2(2\tau, 2z)^{-1}\eta(\tau)\eta(2\tau)$

B.3.12 $\ell = 12$, $X = A_{11}D_7E_6$

We have $m^X = 12$, so the umbral McKay-Thompson series $H_g^{(12)} = (H_{g,r}^{(12)})$ associated to $g \in G^{(12)} \simeq \mathbb{Z}/2\mathbb{Z}$ is a 24-vector-valued function with components indexed by $r \in \mathbb{Z}/24\mathbb{Z}$. We have $H_{g,r}^{(12)} = -H_{g,-r}^{(12)}$, so it suffices to specify the $H_{g,r}^{(12)}$ for $1 \leq r \leq 11$.

To define $H_e^{(12)} = (H_{e,r}^{(12)})$, first define $h(\tau) = (h_r(\tau))$ by requiring that

$$-2\Psi_{1,1}(\tau, z)\left(\varphi_1^{(12)}(\tau, z) + \varphi_2^{(12)}(\tau, z)\right) = -24\mu_{12,0}(\tau, z) + \sum_{r \bmod 24} h_r(\tau)\theta_{12,r}(\tau, z), \tag{B.76}$$

where

$$\begin{aligned}\varphi_1^{(12)}(\tau, z) &:= 3\varphi_1^{(3)}(\tau, z)\varphi_1^{(10)}(\tau, z) - 8\varphi_1^{(4)}(\tau, z)\varphi_1^{(9)}(\tau, z) + \varphi_1^{(5)}(\tau, z)\varphi_1^{(8)}(\tau, z),\\ \varphi_2^{(12)}(\tau, z) &:= 4\varphi_1^{(4)}(\tau, z)\varphi_1^{(9)}(\tau, z) - \varphi_1^{(5)}(\tau, z)\varphi_1^{(8)}(\tau, z) - \varphi_1^{(12)}(\tau, z).\end{aligned} \tag{B.77}$$

(Cf. (B.27), (B.35), (B.39), (B.54), (B.56), (B.62), (B.69).) Now define the $H_{1A,r}^{(12)}$ for $r \neq 0 \mod 3$ by setting

$$\begin{aligned}H_{1A,1}^{(12)}(\tau) &:= \frac{1}{24}\left(3h_1(\tau) + h_7(\tau)\right),\\ H_{1A,2}^{(12)}(\tau), H_{1A,10}^{(12)}(\tau) &:= \frac{1}{24}\left(h_2(\tau) + h_{10}(\tau)\right),\\ H_{1A,4}^{(12)}(\tau), H_{1A,8}^{(12)}(\tau) &:= \frac{1}{12}\left(h_4(\tau) + h_8(\tau)\right),\\ H_{1A,5}^{(12)}(\tau) &:= \frac{1}{24}\left(3h_5(\tau) + h_{11}(\tau)\right),\\ H_{1A,7}^{(12)}(\tau) &:= \frac{1}{24}\left(h_1(\tau) + 3h_7(\tau)\right),\\ H_{1A,11}^{(12)}(\tau) &:= \frac{1}{24}\left(h_5(\tau) + 3h_{11}(\tau)\right).\end{aligned} \tag{B.78}$$

For $r = 0 \mod 3$ set

$$H_{1A,r}^{(12)}(\tau) := \frac{1}{12}h_r(\tau), \tag{B.79}$$

and define $H_{2A,r}^{(12)}$ by requiring

$$H_{2A,r}^{(12)}(\tau) := -(-1)^r H_{1A,r}^{(12)}(\tau). \tag{B.80}$$

B.3.13 $\ell = 12+4$, $X = E_6^4$

We have $m^X = 12$, so the umbral McKay-Thompson series $H_g^{(12+4)} = (H_{g,r}^{(12+4)})$ associated to $g \in G^{(12+4)}$ is a 24-vector-valued function with components indexed by $r \in \mathbb{Z}/24\mathbb{Z}$. In addition to the identity $H_{g,r}^{(12+4)} = -H_{g,-r}^{(12+4)}$, we have $H_{g,r}^{(12+4)} = 0$ for $r \in \{2,3,6,9,10\}$, $H_{g,1}^{(12+4)} = H_{g,7}^{(12+4)}$, $H_{g,4}^{(12+4)} = H_{g,8}^{(12+4)}$, and $H_{g,5}^{(12+4)} = H_{g,11}^{(12+4)}$. Thus it suffices to specify the $H_{g,1}^{(12+4)}$, $H_{g,4}^{(12+4)}$ and $H_{g,5}^{(12+4)}$.

Recall (B.13). Also, set $S_1^{E_6}(\tau) := S_{12,1}(\tau) + S_{12,7}(\tau)$, and $S_5^{E_6}(\tau) := S_{12,5}(\tau) + S_{12,11}(\tau)$. For $r = 1$ define

$$\begin{aligned}
H_{1A,1}^{(12+4)}(\tau) &:= H_{1A,1}^{(12)}(\tau) + H_{1A,7}^{(12)}(\tau),\\
H_{2B,1}^{(12+4)}(\tau) &:= [-\tfrac{1}{48}]\left(H_{8AB,1}^{(6)}(\tau/2) - H_{8AB,5}^{(6)}(\tau/2)\right),\\
H_{4A,1}^{(12+4)}(\tau) &:= \frac{1}{S_1^{E_6}(\tau)^2 - S_5^{E_6}(\tau)^2}\left(-2\frac{\eta(2\tau)^8}{\eta(\tau)^4}S_1^{E_6}(\tau) + 8\frac{\eta(\tau)^4\eta(4\tau)^4}{\eta(2\tau)^4}S_5^{E_6}(\tau)\right),\\
H_{3A,1}^{(12+4)}(\tau) &:= [-\tfrac{1}{48}]\left(H_{3A,1}^{(6)}(\tau/2) - H_{3A,5}^{(6)}(\tau/2)\right),\\
H_{8AB,1}^{(12+4)}(\tau) &:= \frac{1}{S_1^{E_6}(\tau)^2 - S_5^{E_6}(\tau)^2}\left(-2F_{8AB,1}^{(12+4)}(\tau)S_1^{E_6}(\tau) + 12F_{8AB,5}^{(12+4)}(\tau/2)S_5^{E_6}(\tau)\right).
\end{aligned} \tag{B.81}$$

In the expression for $g \in 8AB$, we write $F_{8AB,1}^{(12+4)}$ for the unique modular form of weight 2 for $\Gamma_0(32)$ such that

$$F_{8AB,1}^{(12+4)}(\tau) = 1 + 12q + 4q^2 - 24q^5 - 16q^6 - 8q^8 + O(q^9), \tag{B.82}$$

and we write $F_{8AB,5}^{(12+4)}$ for the unique modular form of weight 2 for $\Gamma_0(64)$ such that

$$F_{8AB,5}^{(12+4)}(\tau) = 3q + 4q^3 + 6q^5 - 8q^7 - 9q^9 + 12q^{11} - 18q^{13} - 24q^{15} + O(q^{17}). \tag{B.83}$$

For $r = 4$ define

$$\begin{aligned}
H_{1A,4}^{(12+4)}(\tau) &:= H_{1A,4}^{(12)}(\tau) + H_{1A,8}^{(12)}(\tau),\\
H_{3A,4}^{(12+4)}(\tau) &:= H_{3A,2}^{(6)}(\tau/2) + H_{3A,4}^{(6)}(\tau/2),
\end{aligned} \tag{B.84}$$

and set $H_{g,4}^{(12+4)}(\tau) := 0$ for $g \in 2B \cup 4A \cup 8AB$.

For $r = 5$ define

$$\begin{aligned}
H_{1A,5}^{(12+4)}(\tau) &:= H_{1A,5}^{(12)}(\tau) + H_{1A,11}^{(12)}(\tau),\\
H_{2B,5}^{(12+4)}(\tau) &:= [-\tfrac{25}{48}]\left(H_{8AB,5}^{(6)}(\tau/2) - H_{8AB,1}^{(6)}(\tau/2)\right),\\
H_{4A,5}^{(12+4)}(\tau) &:= \frac{1}{S_1^{E_6}(\tau)^2 - S_5^{E_6}(\tau)^2}\left(2\frac{\eta(2\tau)^8}{\eta(\tau)^4}S_5^{E_6}(\tau) - 8\frac{\eta(\tau)^4\eta(4\tau)^4}{\eta(2\tau)^4}S_1^{E_6}(\tau)\right),\\
H_{3A,5}^{(12+4)}(\tau) &:= [-\tfrac{25}{48}]\left(H_{3A,5}^{(6)}(\tau/2) - H_{3A,1}^{(6)}(\tau/2)\right),\\
H_{8AB,5}^{(12+4)}(\tau) &:= \frac{1}{S_1^{E_6}(\tau)^2 - S_5^{E_6}(\tau)^2}\left(2F_{8AB,1}^{(12+4)}(\tau)S_5^{E_6}(\tau) - 12F_{8AB,5}^{(12+4)}(\tau/2)S_1^{E_6}(\tau)\right).
\end{aligned} \tag{B.85}$$

Finally, define $H^{(12+4)}_{g,r}$ for $g \in 2A \cup 6A$ by setting

$$\begin{aligned} H^{(12+4)}_{2A,r}(\tau) &:= -(-1)^r H^{(12+4)}_{1A,r}(\tau), \\ H^{(12+4)}_{6A,r}(\tau) &:= -(-1)^r H^{(12+4)}_{3A,r}(\tau). \end{aligned} \tag{B.86}$$

B.3.14 $\ell = 13$, $X = A_{12}^2$

We have $m^X = 13$, so the umbral McKay-Thompson series $H^{(13)}_g = (H^{(13)}_{g,r})$ associated to $g \in G^{(13)} = G^X \simeq \mathbb{Z}/4\mathbb{Z}$ is a 26-vector-valued function, with components indexed by $r \in \mathbb{Z}/26\mathbb{Z}$.

Define $H^{(13)}_g = (H^{(13)}_{g,r})$ for $g \in G^{(13)}$ by requiring that

$$\psi^{(13)}_g(\tau,z) = -\chi^{(13)}_g \mu^0_{13,0}(\tau,z) - \bar{\chi}^{(13)}_g \mu^1_{13,0}(\tau,z) + \sum_{r \mod 26} H^{(13)}_{g,r}(\tau)\theta_{13,r}(\tau,z), \tag{B.87}$$

where $\chi^{(13)}_g := \chi^{X_A}_g$ and $\bar{\chi}^{(13)}_g := \bar{\chi}^{X_A}_g$ (cf. Table 28), and the $\psi^{(13)}_g$ are meromorphic Jacobi forms of weight 1 and index 13 given explicitly in Table 42.

Table 42: Character Values and Meromorphic Jacobi Forms for $\ell = 13$, $X = A_{12}^2$

$[g]$	$\chi^{(13)}_g$	$\bar{\chi}^{(13)}_g$	$\psi^{(13)}_g(\tau,z)$
1A	2	2	$2i\theta_1(\tau,6z)\theta_1(\tau,z)^{-1}\theta_1(\tau,3z)^{-1}\eta(\tau)^3$
2A	-2	2	$-2i\theta_1(\tau,6z)\theta_2(\tau,z)^{-1}\theta_2(\tau,3z)^{-1}\eta(\tau)^3$
4A	0	0	cf. (B.88)

$$\psi^{(13)}_{4AB}(\tau,z) := -i\theta_2(\tau,6z)\frac{\theta_1(\tau,z+\frac14)\theta_1(\tau,3z+\frac14) - \theta_1(\tau,z-\frac14)\theta_1(\tau,3z-\frac14)}{\theta_2(2\tau,2z)\theta_2(2\tau,6z)}\frac{\eta(2\tau)^2}{\eta(\tau)} \tag{B.88}$$

For use later on we note that $\psi^{(13)}_{1A} = -2\Psi_{1,1}\varphi^{(13)}_1$, where

$$\varphi^{(13)}_1(\tau,z) := \frac{\theta_1(\tau,z)\theta_1(\tau,6z)}{\theta_1(\tau,2z)\theta_1(\tau,3z)}. \tag{B.89}$$

B.3.15 $\ell = 14 + 7$, $X = D_8^3$

We have $m^X = 14$, so the umbral McKay-Thompson series $H^{(14+7)}_g = (H^{(14+7)}_{g,r})$ associated to $g \in G^{(14+7)}$ is a 28-vector-valued function with components indexed by $r \in \mathbb{Z}/28\mathbb{Z}$. We have $H^{(14+7)}_{g,r} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H^{(14+7)}_{g,r}$ for r odd. Observing that $H^{(14+7)}_{g,r} = -H^{(14+7)}_{g,-r}$ we may determine $H^{(14+7)}_g$ by requiring that

$$\psi^{(7/2)}_g(\tau,z) = -2\bar{\chi}^{(7/2)}_g i\mu_{7/2,0}(\tau,z) + \sum_{\substack{r\in\mathbb{Z}+1/2 \\ r \mod 7}} e(-r/2)H^{(14+7)}_{g,2r}(\tau)\theta_{7/2,r}(\tau,z), \tag{B.90}$$

where $\bar{\chi}^{(7/2)}_g := \bar{\chi}^{X_D}_g$ is the number of fixed points of $g \in G^{(14+7)} \simeq S_3$ in the defining permutation representation on 3 points. The $\psi^{(7/2)}_g$ are the meromorphic Jacobi forms of weight 1 and index 7/2 defined in Table 43.

Table 43: Character Values and Meromorphic Jacobi Forms for $\ell = 14+7$, $X = D_8^3$

$[g]$	$\bar{\chi}_g^{(7/2)}$	$\psi_g^{(7/2)}(\tau,z)$
1A	3	$2i\theta_1(\tau,3z)\theta_1(\tau,z)^{-2}\eta(\tau)^3$
2A	1	$2i\theta_2(\tau,3z)\theta_1(\tau,z)^{-1}\theta_2(\tau,z)^{-1}\eta(\tau)^3$
3A	0	$-2i\theta_1(\tau,z)\theta_1(\tau,3z)\theta_1(3\tau,3z)^{-1}\eta(3\tau)$

B.3.16 $\ell = 16$, $X = A_{15}D_9$

We have $m^X = 16$, so the umbral McKay-Thompson series $H_g^{(16)} = (H_{g,r}^{(16)})$ associated to $g \in G^{(16)} \simeq \mathbb{Z}/2\mathbb{Z}$ is a 32-vector-valued function with components indexed by $r \in \mathbb{Z}/32\mathbb{Z}$. We have $H_{g,r}^{(16)} = -H_{g,-r}^{(16)}$, so it suffices to specify the $H_{g,r}^{(16)}$ for $1 \leq r \leq 15$.

To define $H_g^{(16)} = (H_{g,r}^{(16)})$ for $g = e$, first define $h(\tau) = (h_r(\tau))$ by requiring that

$$-6\Psi_{1,1}(\tau,z)\left(\varphi_1^{(16)}(\tau,z) + \frac{1}{2}\varphi_2^{(16)}(\tau,z)\right) = -24\mu_{16,0}(\tau,z) + \sum_{r \bmod 32} h_r(\tau)\theta_{16,r}(\tau,z), \tag{B.91}$$

where

$$\begin{aligned}\varphi_1^{(16)}(\tau,z) &:= 8\varphi_1^{(4)}(\tau,z)\varphi_1^{(13)}(\tau,z) - \varphi_1^{(5)}(\tau,z)\varphi_1^{(12)}(\tau,z) + \varphi_1^{(7)}(\tau,z)\varphi_1^{(10)}(\tau,z),\\ \varphi_2^{(16)}(\tau,z) &:= 12\varphi_1^{(4)}(\tau,z)\varphi_1^{(13)}(\tau,z) - \varphi_1^{(5)}(\tau,z)\varphi_1^{(12)}(\tau,z) - 3\varphi_1^{(16)}(\tau,z).\end{aligned} \tag{B.92}$$

(Cf. (B.35), (B.39), (B.54), (B.69), (B.77), (B.89).) Now define the $H_{1A,r}^{(16)}$ by setting

$$H_{1A,r}^{(16)}(\tau) := \frac{1}{12}h_r(\tau) \tag{B.93}$$

for r odd. For r even, $2 \leq r \leq 14$, use

$$H_{1A,r}^{(16)}(\tau) := \frac{1}{24}\left(h_r(\tau) + h_{16-r}(\tau)\right). \tag{B.94}$$

Define $H_{2A,r}^{(16)}$ by requiring

$$H_{2A,r}^{(16)}(\tau) := -(-1)^r H_{1A,r}^{(16)}(\tau). \tag{B.95}$$

B.3.17 $\ell = 18$, $X = A_{17}E_7$

We have $m^X = 18$, so the umbral McKay-Thompson series $H_g^{(18)} = (H_{g,r}^{(18)})$ associated to $g \in G^{(18)} \simeq \mathbb{Z}/2\mathbb{Z}$ is a 36-vector-valued function with components indexed by $r \in \mathbb{Z}/36\mathbb{Z}$. We have $H_{g,r}^{(18)} = -H_{g,-r}^{(18)}$, so it suffices to specify the $H_{g,r}^{(18)}$ for $1 \leq r \leq 17$.

To define $H_g^{(18)} = (H_{g,r}^{(18)})$ for $g = e$, first define $h(\tau) = (h_r(\tau))$ by requiring that

$$-24\Psi_{1,1}(\tau,z)\phi^{(18)}(\tau,z) = -24\mu_{18,0}(\tau,z) + \sum_{r \bmod 36} h_r(\tau)\theta_{18,r}(\tau,z), \tag{B.96}$$

where

$$\phi^{(18)} := \frac{1}{12}\left(\varphi_1^{(18)} + \frac{1}{3}\varphi_3^{(18)} + 4\frac{\theta_1^{12}}{\eta^{12}}\left(\varphi_1^{(12)} + 2\varphi_2^{(12)} + \frac{1}{3}\varphi_3^{(12)}\right)\right). \tag{B.97}$$

For the definition of $\phi^{(18)}$ we require

$$\begin{aligned}
\varphi_2^{(9)}(\tau,z) &:= \varphi_1^{(4)}(\tau,z)\varphi_1^{(6)}(\tau,z) - 4\varphi_1^{(5)}(\tau,z)^2 - 4\varphi_1^{(9)}(\tau,z),\\
\varphi_1^{(11)}(\tau,z) &:= 3\varphi_1^{(5)}(\tau,z)\varphi_1^{(7)}(\tau,z) + 2\varphi_1^{(3)}(\tau,z)\varphi_1^{(9)}(\tau,z) - \varphi_1^{(4)}(\tau,z)\varphi_1^{(8)}(\tau,z),\\
\varphi_3^{(12)}(\tau,z) &:= \varphi_1^{(4)}(\tau,z)\varphi_2^{(9)}(\tau,z),\\
\varphi_1^{(14)}(\tau,z) &:= 3\varphi_1^{(5)}(\tau,z)\varphi_1^{(10)}(\tau,z) + \varphi_1^{(3)}(\tau,z)\varphi_1^{(12)}(\tau,z) - 4\varphi_1^{(4)}(\tau,z)\varphi_1^{(11)}(\tau,z),\\
\varphi_1^{(15)}(\tau,z) &:= \varphi_1^{(5)}(\tau,z)\varphi_1^{(11)}(\tau,z) + 6\varphi_1^{(3)}(\tau,z)\varphi_1^{(13)}(\tau,z) - \varphi_1^{(4)}(\tau,z)\varphi_1^{(12)}(\tau,z),\\
\varphi_2^{(15)}(\tau,z) &:= \varphi_1^{(4)}(\tau,z)\varphi_1^{(12)}(\tau,z) - 2\varphi_1^{(5)}(\tau,z)\varphi_1^{(11)}(\tau,z) - 2\varphi_1^{(15)}(\tau,z),\\
\varphi_1^{(18)}(\tau,z) &:= \varphi_1^{(5)}(\tau,z)\varphi_1^{(14)}(\tau,z) + 3\varphi_1^{(3)}(\tau,z)\varphi_1^{(16)}(\tau,z) - 4\varphi_1^{(4)}(\tau,z)\varphi_1^{(15)}(\tau,z),\\
\varphi_3^{(18)}(\tau,z) &:= \varphi_1^{(4)}(\tau,z)\varphi_2^{(15)}(\tau,z),
\end{aligned} \tag{B.98}$$

in addition to the other $\varphi_k^{(m)}$ that have appeared already. Now define the $H_{1A,r}^{(18)}$ by setting

$$H_{1A,r}^{(18)}(\tau) := \frac{1}{24}h_r(\tau) \tag{B.99}$$

for r even. For r odd, use

$$\begin{aligned}
H_{1A,1}^{(18)}(\tau) &:= \frac{1}{24}\left(2h_1(\tau) + h_{17}(\tau)\right),\\
H_{1A,3}^{(18)}(\tau) &:= \frac{1}{24}\left(h_3(\tau) + h_9(\tau)\right),\\
H_{1A,5}^{(18)}(\tau) &:= \frac{1}{24}\left(2h_5(\tau) + h_{13}(\tau)\right),\\
H_{1A,7}^{(18)}(\tau) &:= \frac{1}{24}\left(2h_7(\tau) + h_{11}(\tau)\right),\\
H_{1A,9}^{(18)}(\tau) &:= \frac{1}{24}\left(h_3(\tau) + 2h_9(\tau) + h_{15}(\tau)\right),\\
H_{1A,11}^{(18)}(\tau) &:= \frac{1}{24}\left(h_7(\tau) + 2h_{11}(\tau)\right),\\
H_{1A,13}^{(18)}(\tau) &:= \frac{1}{24}\left(h_5(\tau) + 2h_{13}(\tau)\right),\\
H_{1A,15}^{(18)}(\tau) &:= \frac{1}{24}\left(h_{15}(\tau) + h_9(\tau)\right),\\
H_{1A,17}^{(18)}(\tau) &:= \frac{1}{24}\left(h_1(\tau) + 2h_{17}(\tau)\right).
\end{aligned} \tag{B.100}$$

Define $H_{2A,r}^{(18)}$ in the usual way for root systems with a type A component, by requiring

$$H_{2A,r}^{(18)}(\tau) := -(-1)^r H_{1A,r}^{(18)}(\tau). \tag{B.101}$$

B.3.18 $\ell = 18+9$, $X = D_{10}E_7^2$

We have $m^X = 18$, so the umbral McKay-Thompson series $H_g^{(18+9)} = (H_{g,r}^{(18+9)})$ associated to $g \in G^{(18+9)} \simeq \mathbb{Z}/2\mathbb{Z}$ is a 36-vector-valued function with components indexed by $r \in \mathbb{Z}/36\mathbb{Z}$. We have $H_{g,r}^{(18+9)} = -H_{g,-r}^{(18+9)}$, $H_{g,r}^{(18+9)} = H_{g,18-r}^{(18+9)}$ for $1 \leq r \leq 17$, and $H_{g,r}^{(18+9)} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H_{g,r}^{(18+9)}$ for $r \in \{1,3,5,7,9\}$.

Define

$$\begin{aligned} H_{1A,r}^{(18+9)}(\tau) &:= H_{1A,r}^{(18)}(\tau) + H_{1A,18-r}^{(18)}(\tau), \\ H_{2A,r}^{(18+9)}(\tau) &:= H_{1A,r}^{(18)}(\tau) - H_{1A,18-r}^{(18)}(\tau), \end{aligned} \tag{B.102}$$

for $r \in \{1,3,5,7,9\}$.

B.3.19 $\ell = 22+11$, $X = D_{12}^2$

We have $m^X = 22$, so the umbral McKay-Thompson series $H_g^{(22+11)} = (H_{g,r}^{(22+11)})$ associated to $g \in G^{(22+11)} \simeq \mathbb{Z}/2\mathbb{Z}$ is a 44-vector-valued function with components indexed by $r \in \mathbb{Z}/44\mathbb{Z}$. We have $H_{g,r}^{(22+11)} = -H_{g,-r}^{(22+11)}$ and $H_{g,r}^{(22+11)} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H_{g,r}^{(22+11)}$ for r odd. Observing that $H_{g,r}^{(22+11)} = -H_{g,-r}^{(22+11)}$ we may determine $H_g^{(22+11)}$ by requiring that

$$\psi_g^{(11/2)}(\tau,z) = -2\bar{\chi}_g^{(11/2)} i\mu_{11/2,0}(\tau,z) + \sum_{\substack{r \in \mathbb{Z}+1/2 \\ r \mod 11}} e(-r/2) H_{g,2r}^{(22+11)}(\tau)\theta_{11/2,r}(\tau,z), \tag{B.103}$$

where $\bar{\chi}_{1A}^{(11/2)} := 2$, $\bar{\chi}_{2A}^{(11/2)} := 0$, and the $\psi_g^{(11/2)}$ are the meromorphic Jacobi forms of weight 1 and index 11/2 defined as follows.

$$\begin{aligned} \psi_{1A}^{(11/2)}(\tau,z) &:= 2i\frac{\theta_1(\tau,4z)}{\theta_1(\tau,z)\theta_1(\tau,2z)}\eta(\tau)^3 \\ \psi_{2A}^{(11/2)}(\tau,z) &:= -2i\frac{\theta_1(\tau,4z)}{\theta_2(\tau,z)\theta_2(\tau,2z)}\eta(\tau)^3 \end{aligned} \tag{B.104}$$

B.3.20 $\ell = 25$, $X = A_{24}$

We have $m^X = 25$, so for $g \in G^{(25)} \simeq \mathbb{Z}/2\mathbb{Z}$, the associated umbral McKay-Thompson series $H_g^{(25)} = (H_{g,r}^{(25)})$ is a 50-vector-valued function, with components indexed by $r \in \mathbb{Z}/50\mathbb{Z}$, satisfying $H_{g,r}^{(25)} = -H_{g,-r}^{(25)}$, and in particular, $H_{g,r}^{(25)} = 0$ for $r = 0 \mod 25$. So it suffices to specify the $H_{g,r}^{(25)}$ for $1 \leq r \leq 24$.

Define $H_g^{(25)} = (H_{g,r}^{(25)})$ for $g = e$ by requiring that

$$-\Psi_{1,1}(\tau,z)\varphi_1^{(25)}(\tau,z) = -\mu_{25,0}(\tau,z) + \sum_{r \mod 50} H_{e,r}^{(25)}(\tau)\theta_{25,r}(\tau,z), \tag{B.105}$$

where

$$\varphi_1^{(25)}(\tau,z) := \frac{1}{2}\varphi_1^{(5)}(\tau,z)\varphi_1^{(21)}(\tau,z) - \varphi_1^{(7)}(\tau,z)\varphi_1^{(19)}(\tau,z) + \frac{1}{2}\varphi_1^{(13)}(\tau,z)^2. \tag{B.106}$$

For the definition of $\varphi_1^{(25)}$ we require

$$\begin{aligned}
\varphi_1^{(17)}(\tau,z) &:= 4\varphi_1^{(5)}(\tau,z)\varphi_1^{(13)}(\tau,z) - \varphi_1^{(9)}(\tau,z)^2,\\
\varphi_1^{(19)}(\tau,z) &:= \varphi_1^{(4)}(\tau,z)\varphi_1^{(16)}(\tau,z) + 2\varphi_1^{(7)}(\tau,z)\varphi_1^{(13)}(\tau,z) - \varphi_1^{(5)}(\tau,z)\varphi_1^{(15)}(\tau,z),\\
\varphi_1^{(21)}(\tau,z) &:= \varphi_1^{(5)}(\tau,z)\varphi_1^{(17)}(\tau,z) - 2\varphi_1^{(9)}(\tau,z)\varphi_1^{(13)}(\tau,z),
\end{aligned} \tag{B.107}$$

in addition to the other $\varphi_k^{(m)}$ that have appeared already. Define $H_{2A,r}^{(25)}$ in the usual way for root systems with a type A component, by requiring

$$H_{2A,r}^{(18)}(\tau) := -(-1)^r H_{1A,r}^{(18)}(\tau). \tag{B.108}$$

B.3.21 $\ell = 30+15$, $X = D_{16}E_8$

We have $m^X = 30$, so the umbral McKay-Thompson series $H_g^{(30+15)} = (H_{g,r}^{(30+15)})$ associated to $g \in G^{(30+15)} = \{e\}$ is a 60-vector-valued function with components indexed by $r \in \mathbb{Z}/60\mathbb{Z}$. We have $H_{e,r}^{(30+15)} = -H_{e,-r}^{(30+15)}$, $H_{e,r}^{(30+15)} = H_{e,30-r}^{(30+15)}$ for $1 \le r \le 29$, and $H_{e,r}^{(30+15)} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H_{e,r}^{(30+15)}$ for $r \in \{1,3,5,7,9,11,13,15\}$.

Define

$$\begin{aligned}
H_{1A,1}^{(30+15)}(\tau) &:= \frac{1}{2}\left(H_{1A,1}^{(30+6,10,15)} + [-\tfrac{1}{120}]H_{3A,1}^{(10+5)}(\tau/3)\right),\\
H_{1A,3}^{(30+15)}(\tau) &:= [-\tfrac{9}{120}]H_{3A,3}^{(10+5)}(\tau/3),\\
H_{1A,5}^{(30+15)}(\tau) &:= [-\tfrac{25}{120}]H_{3A,5}^{(10+5)}(\tau/3),\\
H_{1A,7}^{(30+15)}(\tau) &:= \frac{1}{2}\left(H_{1A,7}^{(30+6,10,15)} + [-\tfrac{49}{120}]H_{3A,3}^{(10+5)}(\tau/3)\right),\\
H_{1A,11}^{(30+15)}(\tau) &:= \frac{1}{2}\left(H_{1A,1}^{(30+6,10,15)} - [-\tfrac{1}{120}]H_{3A,1}^{(10+5)}(\tau/3)\right),\\
H_{1A,13}^{(30+15)}(\tau) &:= \frac{1}{2}\left(H_{1A,7}^{(30+6,10,15)} - [-\tfrac{49}{120}]H_{3A,3}^{(10+5)}(\tau/3)\right),\\
H_{1A,15}^{(30+15)}(\tau) &:= -[-\tfrac{105}{120}]H_{3A,5}^{(10+5)}(\tau/3).
\end{aligned} \tag{B.109}$$

B.3.22 $\ell = 30+6,10,15$, $X = E_8^3$

We have $m^X = 30$, and $G^{(30+6,10,15)} = G^X \simeq S_3$. The umbral McKay-Thompson series $H^{(30+6,10,15)}$ is a 60-vector-valued function with components indexed by $r \in \mathbb{Z}/60\mathbb{Z}$. We have

$$H_{g,r}^{(30+6,10,15)}(\tau) = \begin{cases} \pm H_{g,1}^{(30+6,10,15)} & \text{if } r = \pm1, \pm11, \pm19, \pm29 \mod 60,\\ \pm H_{g,7}^{(30+6,10,15)} & \text{if } r = \pm7, \pm13, \pm17, \pm27 \mod 60,\\ 0 & \text{else,} \end{cases} \tag{B.110}$$

so it suffices to specify the $H^{(30+6,10,15)}_{g,r}$ for $r=1$ and $r=7$. These functions may be defined as follows.

$$\begin{aligned}
H^{(30+6,10,15)}_{1A,1} &:= -2\frac{1}{\eta(\tau)^2}\left(\sum_{k,l,m\geq 0}+\sum_{k,l,m<0}\right)(-1)^{k+l+m}q^{(k^2+l^2+m^2)/2+2(kl+lm+mk)+(k+l+m)/2+3/40}\\
H^{(30+6,10,15)}_{2A,1} &:= -2\frac{1}{\eta(2\tau)}\left(\sum_{k,m\geq 0}-\sum_{k,m<0}\right)(-1)^{k+m}q^{3k^2+m^2/2+4km+(2k+m)/2+3/40}\\
H^{(30+6,10,15)}_{3A,1} &:= -2\frac{\eta(\tau)}{\eta(3\tau)}\sum_{k\in\mathbb{Z}}(-1)^k q^{15k^2/2+3k/2+3/40}\\
H^{(30+6,10,15)}_{1A,7} &= -2\frac{1}{\eta(\tau)^2}\left(\sum_{k,l,m\geq 0}+\sum_{k,l,m<0}\right)(-1)^{k+l+m}q^{(k^2+l^2+m^2)/2+2(kl+lm+mk)+3(k+l+m)/2+27/40}\\
H^{(30+6,10,15)}_{2A,7} &= 2\frac{1}{\eta(2\tau)}\left(\sum_{k,m\geq 0}-\sum_{k,m<0}\right)(-1)^{k+m}q^{3k^2+m^2/2+4km+3(2k+m)/2+27/40}\\
H^{(30+6,10,15)}_{3A,7} &= -2\frac{\eta(\tau)}{\eta(3\tau)}\sum_{k\in\mathbb{Z}}(-1)^k q^{15k^2/2+9k/2+27/40}
\end{aligned} \tag{B.111}$$

B.3.23 $\ell = 46+23$, $X = D_{24}$

We have $m^X = 22$, and $G^{(46+23)} = \{e\}$. The umbral McKay-Thompson series $H^{(46+23)}_e = (H^{(46+23)}_{e,r})$ is a 92-vector-valued function with components indexed by $r \in \mathbb{Z}/92\mathbb{Z}$. We have $H^{(46+23)}_{e,r} = -H^{(46+23)}_{e,-r}$ and $H^{(46+23)}_{e,r} = 0$ for $r = 0 \mod 2$, so it suffices to specify the $H^{(46+23)}_{e,r}$ for r odd. Observing that $H^{(46+23)}_{e,r} = -H^{(46+23)}_{e,-r}$ we may determine $H^{(46+23)}_e$ by requiring that

$$\psi^{(23/2)}_e(\tau,z) = -2i\mu_{23/2,0}(\tau,z) + \sum_{\substack{r\in\mathbb{Z}+1/2\\ r \mod 23}} e(-r/2)H^{(46+23)}_{g,2r}(\tau)\theta_{23/2,r}(\tau,z), \tag{B.112}$$

where $\psi^{(23/2)}_e$ is the meromorphic Jacobi forms of weight 1 and index 23/2 defined by setting

$$\psi^{(23/2)}_e(\tau,z) := 2i\frac{\theta_1(\tau,6z)}{\theta_1(\tau,2z)\theta_1(\tau,3z)}\eta(\tau)^3. \tag{B.113}$$

B.4 Rademacher Sums

Let Γ_∞ denote the subgroup of upper-triangular matrices in $SL_2(\mathbb{Z})$. Given $\alpha\in\mathbb{R}$ and $\gamma\in SL_2(\mathbb{Z})$, define $\mathrm{r}^{[\alpha]}_{1/2}(\gamma,\tau) := 1$ if $\gamma\in\Gamma_\infty$. Otherwise, set

$$\mathrm{r}^{[\alpha]}_{1/2}(\gamma,\tau) := e(-\alpha(\gamma\tau-\gamma\infty))\sum_{k\geq 0}\frac{(2\pi i\alpha(\gamma\tau-\gamma\infty))^{n+1/2}}{\Gamma(n+3/2)}, \tag{B.114}$$

where $e(x) := e^{2\pi i x}$. Let n be a positive integer, and suppose that ν is a multiplier system for vector-valued modular forms of weight 1/2 on $\Gamma = \Gamma_0(n)$. Assume that $\nu = (\nu_{ij})$ satisfies $\nu_{11}(T) = e^{\pi i/2m}$, for some basis $\{\mathfrak{e}_i\}$, for some positive integer m, where $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. To this data, attach the Rademacher sum

$$R_{\Gamma,\nu}(\tau) := \lim_{K\to\infty} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_{K,K^2}} \nu(\gamma) e\left(-\frac{\gamma\tau}{4m}\right) \mathfrak{e}_1 \, \mathrm{j}(\gamma,\tau)^{1/2} \, \mathrm{r}_{1/2}^{[-1/4m]}(\gamma,\tau), \tag{B.115}$$

where $\Gamma_{K,K^2} := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma \mid 0 \le c < K, |d| < K^2 \right\}$, and $\mathrm{j}(\gamma,\tau) := (c\tau + d)^{-1}$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. If the expression (B.115) converges then it defines a mock modular form of weight 1/2 for Γ whose shadow is given by an explicitly identifiable Poincaré series. We refer to [13] for a review of this, and to [53] for a more general and detailed discussion.

Convergence of (B.115) can be shown by rewriting the Fourier expansion as in [53, Theorem 2] in terms of a sum of Kloostermann sums weighted by Bessel functions. This expression converges at $w = 1/2$ by the analysis discussed at the end of §3, following the method of Hooley as adapted by Gannon. That analysis requires not only establishing that the expressions converge, but also explicitly bounding the rates of convergence.

For the special case that $X = A_8^3$ we require 8-vector-valued functions $\check{t}_g^{(9)} = (\check{t}_{g,r}^{(9)})$ for $g \in G^X$ with order 3 or 6. For such g, define $\check{t}_{g,r}^{(9)}$, for $0 < r < 9$, by setting

$$\check{t}_{3A,r}^{(9)}(\tau) := \begin{cases} 0, & \text{if } r \neq 0 \mod 3, \\ -\theta_{3,3}(\tau,0), & \text{if } r = 3, \\ \theta_{3,0}(\tau,0), & \text{if } r = 6, \end{cases} \tag{B.116}$$

in the case that g has order 3, and

$$\check{t}_{6A,r}^{(9)}(\tau) := \begin{cases} 0, & \text{if } r \neq 0 \mod 3, \\ -\theta_{3,3}(\tau,0), & \text{if } r = 3, \\ -\theta_{3,0}(\tau,0), & \text{if } r = 6, \end{cases} \tag{B.117}$$

when $o(g) = 6$. Here $\theta_{m,r}(\tau,z)$ is as defined in (B.5).

The following result is proved in [17], using an analysis of representations of the metaplectic double cover of $SL_2(\mathbb{Z})$.

Theorem B.1 ([17])**.** *Let X be a Niemeier root system and let $g \in G^X$. Assume that the Rademacher sum $R^X_{\Gamma_0(n_g),\check{\nu}_g^X}$ converges. If $X \neq A_8^3$, or if $X = A_8^3$ and $g \in G^X$ does not satisfy $o(g) = 0 \mod 3$, then we have*

$$\check{H}_g^X(\tau) = -2R^X_{\Gamma_0(n_g),\check{\nu}_g^X}. \tag{B.118}$$

If $X = A_8^3$ and $g \in G^X$ satisfies $o(g) = 0 \mod 3$ then

$$\check{H}_{g,r}^X(\tau) = -2R^X_{\Gamma_0(n_g),\check{\nu}_g^X}(\tau) + \check{t}_g^{(9)}(\tau). \tag{B.119}$$

The $X = A_1^{24}$ case of Theorem B.1 was proven first in [12], via different methods.

References

[1] D. Alexander, C. Cummins, J. McKay, and C. Simons, *Completely replicable functions*, Groups, combinatorics & geometry (Durham, 1990), London Math. Soc. Lecture Note Ser., vol. 165, Cambridge Univ. Press, Cambridge, 1992, 87–98. MR 1200252 (94g:11029)

[2] R. Borcherds, *Vertex algebras, Kac-Moody algebras, and the Monster*, Proceedings of the National Academy of Sciences, U.S.A. **83** (1986), no. 10, 3068–3071.

[3] ______, *Generalized Kac-Moody algebras*, J. Algebra **115** (1988), no. 2, 501–512. MR 943273 (89g:17004)

[4] ______, *Central extensions of generalized Kac-Moody algebras*, J. Algebra **140** (1991), no. 2, 330–335. MR 1120425 (92g:17031)

[5] ______, *Monstrous moonshine and monstrous Lie superalgebras*, Invent. Math. **109** (1992), no. 2, 405–444.

[6] ______, *Automorphic forms with singularities on Grassmannians*, Invent. Math. **132** (1998), no. 3, 491–562. MR 1625724 (99c:11049)

[7] K. Bringmann and K. Ono, *The $f(q)$ mock theta function conjecture and partition ranks*, Invent. Math. **165** (2006), no. 2, 243–266. MR 2231957 (2007e:11127)

[8] ______, *Coefficients of harmonic maass forms*, Partitions, q-Series, and Modular Forms (Krishnaswami Alladi and Frank Garvan, eds.), Developments in Mathematics, vol. 23, Springer New York, 2012, pp. 23–38.

[9] J. H. Bruinier and J. Funke, *On two geometric theta lifts*, Duke Math. J. **125** (2004), no. 1, 45–90. MR 2097357 (2005m:11089)

[10] M. C. Cheng, *K3 surfaces, $N = 4$ dyons and the Mathieu group M_{24}*, Commun. Number Theory Phys. **4** (2010), no. 4, 623–657. MR 2793423 (2012e:11076)

[11] M. Cheng and J. Duncan, *The largest Mathieu group and (mock) automorphic forms*, String-Math 2011, Proc. Sympos. Pure Math., vol. 85, Amer. Math. Soc., Providence, RI, 2012, pp. 53–82. MR 2985326

[12] ______, *On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine*, Commun. Number Theory Phys. **6** (2012), no. 3.

[13] ______, *Rademacher Sums and Rademacher Series*, Conformal Field Theory, Automorphic Forms and Related Topics, 2014, pp.143-182, Contributions in Mathematical and Computational Sciences.

[14] ______, *Meromorphic Jacobi Forms of Half-Integral Index and Umbral Moonshine Modules*, in preparation.

[15] M. C. N. Cheng, J. F. R. Duncan, and J. A. Harvey, *Umbral Moonshine*, Commun. Number Theory Phys. **8** (2014), no. 2.

[16] ______, *Umbral Moonshine and the Niemeier Lattices*, Research in the Mathematical Sciences **1** (2014), no. 3.

[17] ______, *Weight One Jacobi Forms and Umbral Moonshine*, in preparation.

[18] J. H. Conway and S. P. Norton, *Monstrous moonshine*, Bull. London Math. Soc. **11** (1979), no. 3, 308–339. MR 554399 (81j:20028)

[19] J. H. Conway and N. J. A. Sloane, *On the enumeration of lattices of determinant one*, J. Number Theory **15** (1982), no. 1, 83–94. MR 666350 (84b:10047)

[20] ______, *Sphere packings, lattices and groups*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 290, Springer-Verlag, New York, 1999, With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov. MR 1662447 (2000b:11077)

[21] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker, and R.A. Wilson, *Atlas of finite groups. Maximal subgroups and ordinary characters for simple groups. With comput. assist. from J. G. Thackray.*, Oxford: Clarendon Press, 1985.

[22] A. Dabholkar, S. Murthy and D. Zagier, *Quantum Black Holes, Wall Crossing, and Mock Modular Forms*, arXiv:1208.4074.

[23] L. Dixon, P. Ginsparg and J. Harvey, *Beauty and the beast: superconformal symmetry in a Monster module*, Comm. Math. Phys. **119** (1998), no. 2, 221–241. MR 968697 (90b:81119)

[24] J. Duncan, M. Griffin, and K. Ono, *Moonshine*, Research in the Mathematical Sciences (2015) 2:11.

[25] J. Duncan and J. Harvey, *The Umbral Mooonshine Module for the Unique Unimodular Niemeier Root System*, arXiv:1412.8191.

[26] J. Duncan and A. O'Desky, *Super Vertex Algebras, Meromorphic Jacobi Forms, and Umbral Moonshine*, in preparation.

[27] T. Eguchi and K. Hikami, *Note on Twisted Elliptic Genus of K3 Surface*, Phys.Lett. **B694** (2011), 446–455.

[28] T. Eguchi, H. Ooguri, and Y. Tachikawa, *Notes on the K3 Surface and the Mathieu group* M_{24}, Exper.Math. **20** (2011), 91–96.

[29] M. Eichler and D. Zagier, *The theory of Jacobi forms*, Birkhäuser, 1985.

[30] I. Frenkel, J. Lepowsky, and A. Meurman, *A natural representation of the Fischer-Griess Monster with the modular function J as character*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 10, Phys. Sci., 3256–3260. MR MR747596 (85e:20018)

[31] ______, *A moonshine module for the Monster*, Vertex operators in mathematics and physics (Berkeley, Calif., 1983), Math. Sci. Res. Inst. Publ., vol. 3, Springer, New York, 1985, pp. 231–273. MR 86m:20024

[32] ———, *Vertex operator algebras and the Monster*, Pure and Applied Mathematics, vol. 134, Academic Press Inc., Boston, MA, 1988. MR 90h:17026

[33] M. Gaberdiel, S. Hohenegger, and R. Volpato, *Mathieu Moonshine in the elliptic genus of K3*, JHEP **1010** (2010), 062.

[34] ———, *Mathieu twining characters for K3*, JHEP **1009** (2010), 058, 19 pages.

[35] T. Gannon, *Monstrous moonshine: the first twenty-five years*, Bull. London Math. Soc. **38** (2006), no. 1, 1–33. MR 2201600 (2006k:11071)

[36] ———, *Moonshine beyond the Monster*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2006, The bridge connecting algebra, modular forms and physics. MR 2257727 (2008a:17032)

[37] ———, *Much ado about Mathieu*, arXiv:1211.5531.

[38] C. Hooley, *On the number of divisors of a quadratic polynomial*, Acta Math. **110** (1963), 97-114. MR 0153648

[39] J. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag, New York, 1972, Graduate Texts in Mathematics, Vol. 9. MR 0323842 (48 #2197)

[40] W. Kohnen, *Fourier coefficients of half-integral weight*, Math. Ann. **271** (1985), 237–268. MR 0783554

[41] Ö. Imamoḡlu, M. Raum, and O. Richer, *Holomorphic projections and Ramanujan's mock theta functions*, Proc. Natl. Acad. Sci., USA, **111** (2014), 3961–3967.

[42] M. Mertens, *Eichler-Selberg type identities for mixed mock modular forms*, arXiv:1404.5491.

[43] H. V. Niemeier, *Definite quadratische Formen der Dimension* 24 *und Diskriminante* 1, J. Number Theory **5** (1973), 142–178. MR 0316384 (47 #4931)

[44] K. Ono, *The web of modularity: arithmetic of the coefficients of modular forms and q-series*, CBMS Regional Conf. Ser. No. 102, Amer. Math. Soc., 2004, Providence MR 2020489 (2005c:11053)

[45] K. Ono, *Unearthing the visions of a master: harmonic Maass forms and number theory*, Current developments in mathematics, 2008, Int. Press, Somerville, MA, 2009, pp. 347–454. MR 2555930 (2010m:11060)

[46] S. Ramanujan, *The lost notebook and other unpublished papers*, Springer-Verlag, Berlin, 1988, With an introduction by George E. Andrews. MR 947735 (89j:01078)

[47] W. A. Stein et al., *Sage Mathematics Software (Version 6.5)*, The Sage Development Team, 2015, `http://www.sagemath.org`.

[48] J.-P. Serre and H. M. Stark, *Modular forms of weight* 1/2, in Modular functions of one variable, VI (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976), 27–67. Lecture Notes in Math., Vol. 627. Springer, Berlin, 1977.

[49] J. Sturm, *On the congruence of modular forms*, Springer Lect. Notes Math. **1240** (1984), 275-280. MR 0894516 (88h:11031)

[50] J. G. Thompson, *Finite groups and modular functions*, Bull. London Math. Soc. **11** (1979), no. 3, 347–351. MR 554401 (81j:20029)

[51] ______, *Some numerology between the Fischer-Griess Monster and the elliptic modular function*, Bull. London Math. Soc. **11** (1979), no. 3, 352–353. MR MR554402 (81j:20030)

[52] B. B. Venkov, *On the classification of integral even unimodular* 24*-dimensional quadratic forms*, Trudy Mat. Inst. Steklov. **148** (1978), 65–76, 273, Algebra, number theory and their applications. MR 558941 (81d:10024)

[53] D. Whalen, *Vector-Valued Rademacher Sums and Automorphic Integrals*, arXiv:1406.0571.

[54] D. Zagier, *Ramanujan's mock theta functions and their applications (after Zwegers and Ono-Bringmann)*, Astérisque (2009), no. 326, Exp. No. 986, vii–viii, 143–164 (2010), Séminaire Bourbaki. Vol. 2007/2008. MR 2605321 (2011h:11049)

[55] S. Zwegers, *Mock Theta Functions*, Ph.D. thesis, Utrecht University, 2002. arXiv:0807.4834.

J. F. R. Duncan, Dept. of Mathematics & Computer Science, Emory University, Atlanta, GA 30322
E-mail address, J. F. R. Duncan: `john.duncan@emory.edu`

M. J. Griffin, Dept. of Mathematics, Princeton University, Princeton, NJ 08544
E-mail address, M. J. Griffin: `mjg4@princeton.edu`

K. Ono, Dept. of Mathematics & Computer Science, Emory University, Atlanta, GA 30322
E-mail address, K. Ono: `ono@mathcs.emory.edu`

arXiv:1204.2779v3 [math.RT] 17 Oct 2013

Umbral Moonshine[*]

Miranda C. N. Cheng[†]
John F. R. Duncan[‡]
Jeffrey A. Harvey[§]

2013 October 13

Abstract

We describe surprising relationships between automorphic forms of various kinds, imaginary quadratic number fields and a certain system of six finite groups that are parameterised naturally by the divisors of twelve. The Mathieu group correspondence recently discovered by Eguchi–Ooguri–Tachikawa is recovered as a special case. We introduce a notion of extremal Jacobi form and prove that it characterises the Jacobi forms arising by establishing a connection to critical values of Dirichlet series attached to modular forms of weight two. These extremal Jacobi forms are closely related to certain vector-valued mock modular forms studied recently by Dabholkar–Murthy–Zagier in connection with the physics of quantum black holes in string theory. In a manner similar to monstrous moonshine the automorphic forms we identify constitute evidence for the existence of infinite-dimensional graded modules for the six groups in our system. We formulate an umbral moonshine conjecture that is in direct analogy with the monstrous moonshine conjecture of Conway–Norton. Curiously, we find a number of Ramanujan's mock theta functions appearing as McKay–Thompson series. A new feature not apparent in the monstrous case is a property which allows us to predict the fields of definition of certain homogeneous submodules for the groups involved. For four of the groups in our system we find analogues of both the classical McKay correspondence and McKay's monstrous Dynkin diagram observation manifesting simultaneously and compatibly.

---

[*] *MSC2010:* 11F22, 11F37, 11F46, 11F50, 20C34, 20C35

[†] Université Paris 7, UMR CNRS 7586 and LPTHE, Université Paris 6, Paris, France.
E-mail: `chengm@math.jussieu.fr`
[‡] Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106, U.S.A.
E-mail: `john.duncan@case.edu`
[§] Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, IL 60637, U.S.A.
E-mail: `j-harvey@uchicago.edu`

Contents

1 Introduction 4

2 Automorphic Forms 9

2.1 Mock Modular Forms . . . 10
2.2 Jacobi Forms . . . 13
2.3 Meromorphic Jacobi Forms . . . 15
2.4 Superconformal Algebra . . . 17
2.5 Extremal Jacobi Forms . . . 18
2.6 Siegel Modular Forms . . . 30

3 Finite Groups 33

3.1 Specification . . . 34
3.2 Signed Permutations . . . 36
3.3 Realisation Part I . . . 38
3.4 Realisation Part II . . . 41
3.5 Dynkin Diagrams Part I . . . 44
3.6 Dynkin Diagrams Part II . . . 47

4 McKay–Thompson Series 51

4.1 Forms of Higher Level . . . 52
4.2 Lambency Two . . . 56
4.3 Lambency Three . . . 57
4.4 Lambency Four . . . 59
4.5 Lambency Five . . . 61
4.6 Lambencies Seven and Thirteen . . . 62
4.7 Mock Theta Functions . . . 63
4.8 Automorphy . . . 64

5 Conjectures 67

5.1 Modules . . . 68
5.2 Moonshine . . . 69
5.3 Modularity . . . 71
5.4 Discriminants . . . 71
5.5 Geometry and Physics . . . 74

A Modular Forms **78**
A.1 Dedekind Eta Function 78
A.2 Jacobi Theta Functions 78
A.3 Higher Level Modular Forms 79

B Characters **80**
B.1 Irreducible Characters 81
B.2 Euler Characters 85

C Coefficients **88**
C.1 Lambency Two 89
C.2 Lambency Three 90
C.3 Lambency Four 92
C.4 Lambency Five 95
C.5 Lambency Seven 99
C.6 Lambency Thirteen 102

D Decompositions **106**
D.1 Lambency Two 107
D.2 Lambency Three 108
D.3 Lambency Four 109
D.4 Lambency Five 110
D.5 Lambency Seven 111
D.6 Lambency Thirteen 112

References **114**

1 Introduction

The term *monstrous moonshine* was coined by Conway [1] in order to describe the unexpected and mysterious connections between the representation theory of the largest sporadic group—the *Fischer–Griess monster*, $\mathbb{M}$—and modular functions that stemmed from McKay's observation that $196883 + 1 = 196884$, where the summands on the left are degrees of irreducible representations of $\mathbb{M}$ and the number on the right is the coefficient of q in the Fourier expansion of the *elliptic modular invariant*

$$J(\tau) = \sum_{m \geq -1} a(m) q^m = q^{-1} + 196884q + 21493760q^2 + 864299970q^3 + \cdots . \tag{1.1}$$

Thompson expanded upon McKay's observation in [2] and conjectured the existence of an infinite-dimensional monster module

$$V = \bigoplus_{m \geq -1} V_m \tag{1.2}$$

with $\dim V_m = a(m)$ for all m. He also proposed [3] to consider the series, now known as *McKay–Thompson series*, given by

$$T_g(\tau) = \sum_{m \geq -1} \mathrm{tr}_{V_m}(g)\, q^m \tag{1.3}$$

for $g \in \mathbb{M}$, and detailed explorations [1] by Conway–Norton led to the astonishing *moonshine conjecture*:

> For each $g \in \mathbb{M}$ the function T_g is a principal modulus for some genus zero group Γ_g.

(A discrete group $\Gamma < PSL_2(\mathbb{R})$ is said to have *genus zero* if the Riemann surface $\Gamma \backslash \mathbb{H}$ is isomorphic to the Riemann sphere minus finitely many points, and a holomorphic function f on $\mathbb{H}$ is called a *principal modulus* for a genus zero group Γ if it generates the field of Γ-invariant functions on $\mathbb{H}$.)

Thompson's conjecture was verified by Atkin, Fong and Smith (cf. [4, 5]). A more constructive verification was obtained by Frenkel–Lepowsky–Meurman [6, 7] with the explicit construction of a monster module $V = V^\natural$ with graded dimension given by the Fourier expansion (1.1) of the elliptic modular invariant. They used *vertex operators*—structures originating in the dual resonance theory of particle physics and finding contemporaneous application [8, 9] to affine Lie algebras—to recover the non-associative *Griess algebra* structure (developed in the first proof [10] of the existence of the monster) from a subspace of $V^\natural$. Borcherds found a way to attach vertex operators to every element of $V^\natural$ and determined the precise sense in which these operators could be given a commutative associative composition law, and thus arrived at the notion of *vertex algebra* [11], an axiomatisation of the operator product expansion of chiral conformal field theory. The closely related notion of *vertex operator algebra (VOA)* was subsequently introduced by Frenkel–Lepowsky–Meurman [12] and they established that the monster is precisely the group of automorphisms of a VOA structure on $V^\natural$; the Frenkel–Lepowsky–Meurman construction of $V^\natural$ would ultimately prove to furnish the first example of an *orbifold conformal field theory*.

Borcherds introduced the notion of *generalised Kac–Moody algebra* in [13] and by using the VOA structure on $V^\natural$ was able to construct a particular example—the monster Lie algebra—and use the corresponding equivariant denominator identities to arrive at a proof [14] of the Conway–Norton moonshine conjectures. Thus by 1992 monstrous moonshine had already become a phenomenon encompassing elements of finite group theory, modular forms, vertex algebras and generalised Kac–Moody algebras, as well as aspects of conformal field theory and string theory.

Recently Eguchi–Ooguri–Tachikawa have presented evidence [15] for a new kind of moonshine involving the elliptic genus of $K3$ surfaces (the elliptic genus is a topological invariant and therefore independent of the choice of $K3$ surface) and the *largest Mathieu group* M_{24} (cf. §3.1). This connection between $K3$ surfaces and M_{24} becomes apparent only after decomposing the elliptic genus into characters of the $N = 4$ superconformal algebra (cf. [16, 17, 18]). This decomposition process (cf. §2.4) reveals the presence of a mock modular form of weight 1/2 (cf. §2.1) satisfying

$$H^{(2)}(\tau) = \sum_{n=0}^{\infty} c_1^{(2)}(n-1/8)q^{n-1/8} = 2q^{-1/8}(-1+45q+231q^2+770q^3+2277q^4+\cdots) \quad (1.4)$$

and one recognises here the dimensions of several irreducible representations of M_{24} (cf. Table 8).

One is soon led to follow the path forged by Thompson in the case of the monster: to suspect the existence of a graded infinite-dimensional M_{24}-module

$$K^{(2)} = \bigoplus_{n=0}^{\infty} K^{(2)}_{n-1/8} \quad (1.5)$$

with $\dim K^{(2)}_{n-1/8} = c^{(2)}(n-1/8)$ for $n \geq 1$, and to study the analogues $H_g^{(2)}$ of the monstrous McKay–Thompson series obtained by replacing $c^{(2)}(n-1/8) = \dim K^{(2)}_{n-1/8}$ with $\mathrm{tr}_{K^{(2)}_{n-1/8}}(g)$ in (1.4). This idea has been implemented successfully in [19, 20, 21, 22] and provides strong evidence for the existence of such an M_{24}-module $K^{(2)}$. A proof of the existence of $K^{(2)}$ has now been established in [23] although no explicit construction is yet known. In particular there is as yet no known analogue of the vertex operator algebra structure which conjecturally characterises [12] the monster module $V^\natural$.

The strong evidence in support of the M_{24} analogue of Thompson's conjecture invites us to consider the M_{24} analogue of the Conway–Norton moonshine conjectures—this will justify the use of the term moonshine in the M_{24} setting—except that it is not immediately obvious what the analogue should be. Whilst the McKay–Thompson series $H_g^{(2)}$ is a mock modular form of weight $1/2$ on some $\Gamma_g < SL_2(\mathbb{Z})$ for every g in M_{24} [22], it is not the case that Γ_g is a genus zero group for every g, and even if it were, there is no obvious sense in which one mock modular form of weight $1/2$ on some group can "generate" all the others, and thus no obvious analogue of the principal modulus property.

A solution to this problem—the formulation of the moonshine conjecture for M_{24}—was found in [24] (see also §5.2) via an extension of the program that was initiated in [25]; the antecedents of which include Rademacher's pioneering work [26] on the elliptic modular invariant $J(\tau)$, quantum gravity in three dimensions [27, 28, 29], the AdS/CFT correspondence in physics [30, 31, 32], and the application of *Rademacher sums* to these and other settings in string theory [33, 34, 35, 36, 37, 38, 39] (and in particular [40]). To explain the formulation of the moonshine conjecture for M_{24} we recall that in [25] a Rademacher sum $R_\Gamma(\tau)$ is defined for each discrete group $\Gamma < PSL_2(\mathbb{R})$ commensurable with the modular group in such a way as to naturally generalise Rademacher's Poincaré series-like expression for the elliptic modular invariant derived in [26]. It is then shown in [25] that a holomorphic function on the upper-half plane (with invariance group commensurable with $PSL_2(\mathbb{Z})$) is the principal modulus for its invariance group if and only if it coincides with the Rademacher sum attached to this group. Thus the genus property of monstrous moonshine may be reformulated:

> For each g in $\mathbb{M}$ we have $T_g = R_{\Gamma_g}$ where Γ_g is the invariance group of T_g.

Write $R_\Gamma^{(2)}$ to indicate a weight $1/2$ generalisation of the (weight 0) Rademacher sum construction R_Γ studied in [25]. (Note that a choice of multiplier system on Γ is also required.) Then a natural M_{24}-analogue of the Conway–Norton moonshine conjecture comes into view:

> For each g in M_{24} we have $H_g^{(2)} = R_{\Gamma_g}^{(2)}$ where Γ_g is the invariance group of $H_g^{(2)}$.

This statement is confirmed in [24] for the functions $H_g^{(2)}$ that are, at this point, conjecturally attached to M_{24} via the conjectural M_{24}-module $K^{(2)}$.

Through these results we come to envisage the possibility that both monstrous moonshine and the M_{24} observation of Eguchi–Ooguri–Tachikawa will eventually be understood as aspects of one underlying structure which will include finite groups, various kinds of automorphic forms, extended algebras and string theory, and will quite possibly be formulated in terms of the AdS/CFT correspondence in the context of some higher-dimensional string or gravitational theory.

In fact we can expect this *moonshine structure* to encompass more groups beyond the monster and M_{24}: In this paper we identify the M_{24} observation as one of a family of correspondences between finite groups and (vector-valued) mock modular forms, with each member in the family admitting a natural analogue of the Conway–Norton moonshine conjecture according to the philosophy of [25, 24] (see also [41]). For each ℓ in $\Lambda = \{2,3,4,5,7,13\}$—the set of positive integers ℓ such that $\ell - 1$ divides 12—we identify a distinguished Jacobi form $Z^{(\ell)}$, a finite group $G^{(\ell)}$ and a family of vector-valued mock modular forms $H_g^{(\ell)}$ for $g \in G^{(\ell)}$. The Jacobi forms $Z^{(\ell)}$ satisfy an *extremal* condition formulated in terms of unitary irreducible characters of the $N = 4$ superconformal algebra (cf. §2.5), the mock modular form $H^{(\ell)} = H_e^{(\ell)}$ is related to $Z^{(\ell)}$ as $H^{(2)}$ is to the elliptic genus of a $K3$ surface (cf. §§2.4,2.5), the Fourier coefficients of the *McKay–Thompson series* $H_g^{(\ell)}$ support the existence of an infinite-dimensional graded module $K^{(\ell)}$ for $G^{(\ell)}$ playing a rôle analogous to that of $K^{(2)}$ for $G^{(2)} \simeq M_{24}$ (cf. §5.1), and the following *umbral moonshine conjecture* is predicted to hold where $R_\Gamma^{(\ell)}$ is an $(\ell-1)$-vector-valued generalisation (cf. §5.2) of the Rademacher sum construction $R_\Gamma^{(2)}$ studied in [24].

> For each g in $G^{(\ell)}$ we have $H_g^{(\ell)} = R_{\Gamma_g}^{(\ell)}$ where Γ_g is the invariance group of $H_g^{(\ell)}$.

In addition to the above properties with monstrous analogues we find the following new *discriminant property*: The exponents of the powers of q having non-vanishing coefficient in the Fourier development of $H^{(\ell)}$ determine certain imaginary quadratic number fields and predict the existence of dual pairs of irreducible representations of $G^{(\ell)}$ that are defined over these fields and irreducible over $\mathbb{C}$. Moreover, these dual pairs consistently appear as irreducible constituents in homogeneous $G^{(\ell)}$-submodules of $K^{(\ell)}$ in such a way that the degree of the submodule determines the discriminant of the corresponding quadratic field (cf. §5.4).

A main result of this paper is Theorem 2.2, which states that the extremal condition formulated in §2.5 characterises the Jacobi forms $Z^{(\ell)}$ for $\ell \in \{2,3,4,5,7,13\}$. To achieve this we establish a result of independent interest, which also serves to illustrate the depth of the characterisation problem: We show in Theorem 2.3 that the existence of an extremal Jacobi of index $m-1$ implies the vanishing of $L(f,1)$ for all new forms f of weight 2 and level m, where $L(f,s)$ is the Dirichlet series naturally attached to f. According to the Birch–Swinnerton-Dyer conjecture the vanishing of $L(f,1)$ implies that the elliptic curve E_f, attached to f by Eichler–Shimura theory, has rational points of infinite order. Thus it is extremely unexpected that an extremal Jacobi form can exist for all but finitely many values of m, and using estimates [42] due to Ellenberg together with some explicit computations we are able to verify that the only possible values are those for which no non-zero weight 2 forms exist; i.e. $\Gamma_0(m)$ has genus zero. The last step in our proof of Theorem 2.2 is to check the finitely many corresponding finite dimensional spaces of Jacobi forms, for which useful bases have been determined by Gritsenko in [43]. In this we way obtain that the $Z^{(\ell)}$ for $\ell \in \{2,3,4,5,7,13\}$ are precisely the unique, up to scale, weak Jacobi forms of weight zero satisfying the extremal condition (2.37).

In contrast to the monstrous case the McKay–Thompson series $H_g^{(\ell)}$ arising here are mock modular forms, and these are typically not in fact modular but become so after *completion* with respect to a *shadow* function $S_g^{(\ell)}$. It turns out that all the McKay–Thompson series $H_g^{(\ell)}$ for fixed $\ell \in \Lambda$ have shadows that are (essentially) proportional to a single vector-valued *unary theta function* $S^{(\ell)}$ (cf. §2.2) and so it is in a sense the six *moonlight shadows* $S^{(\ell)}$ for $\ell \in \Lambda$ that provide the irreducible information required to uncover the structure that we reveal in this paper. We therefore refer to the phenomena investigated here as *umbral moonshine*.

According to the Oxford English Dictionary, a flame or light that is *lambent* is playing "lightly upon or gliding over a surface without burning it, like a 'tongue of fire'; shining with a soft clear light and without fierce heat." And since the light of umbral moonshine is apparently of this nature, we call the six values in $\Lambda = \{2,3,4,5,7,13\}$ *lambent*, and we refer to the index $\ell \in \Lambda$ as the *lambency* of the connections relating the *umbral group* $G^{(\ell)}$ to the *umbral forms* $Z^{(\ell)}$ and $H_g^{(\ell)}$.

The rest of this paper is organised as follows. In §2 we discuss properties of Jacobi forms, Siegel forms, mock modular forms and mock theta functions. We explain two closely related ways in which Jacobi forms determine mock modular forms, one involving the decomposition into characters of the $N = 4$ superconformal algebra and the other involving a decomposition of meromorphic Jacobi forms into mock modular forms following [44] and [45]. In §2.5 we introduce the Jacobi forms $Z^{(\ell)}$ of weight 0 and index $\ell - 1$ for lambent ℓ and their associated vector-valued mock modular forms $H^{(\ell)}$. We prove (Theorem 2.2) that these functions are characterized by the extremal property (2.37) and we establish the connection (Theorem 2.6) to critical values of automorphic L-functions. We note that the coefficients in the q-expansions of the mock modular forms $H^{(\ell)}$ appear to be connected to the dimensions of irreducible representations of groups $G^{(\ell)}$ which we introduce and study in §3. In §§3.5,3.6 we discuss the remarkable fact that some of these groups manifest both the McKay correspondence relating ADE Dynkin diagrams to finite subgroups of $SU(2)$ as well as a generalisation of his monstrous E_8 observation. In §4 we discuss analytic properties of the McKay–Thompson series $H_g^{(\ell)}$ which are obtained by twisting the mock modular forms $H^{(\ell)}$ by elements $g \in G^{(\ell)}$. We determine the proposed McKay–Thompson series

$H_g^{(\ell)} = \big(H_{g,r}^{(\ell)}\big)$ precisely in terms of modular forms of weight 2 for all but a few g occurring for $\ell \in \{7, 13\}$, and we find that we can identify many of the component functions $H_{g,r}^{(\ell)}$ either with ratios of products of eta functions or with classical mock theta functions introduced by Ramanujan (and others). In §5 we collect our observations into a set of conjectures. These include the analogue of Thompson's conjecture (cf. §5.1), the umbral counterpart to the Conway–Norton moonshine conjecture (cf. §5.2), and a precise formulation of the discriminant property mentioned above (cf. §5.4). In §5.5 we mention some possible connections between our results, the geometry of complex surfaces, and string theory.

Our conventions for modular forms appear in §A, the character tables of the umbral groups $G^{(\ell)}$ appear in §B, tables of Fourier coefficients of low degree for all the proposed McKay–Thompson series $H_g^{(\ell)}$ appear in §C, and tables describing the $G^{(\ell)}$-module structures implied (for low degree) by the $H_g^{(\ell)}$ are collected in §D.

It is important to mention that much of the data presented in the tables of §C was first derived using certain vector-valued generalisations of the Rademacher sum construction that was applied to the functions of monstrous moonshine in [25], and in [24] to the functions attached to M_{24} via the observation of Eguchi–Ooguri–Tachikawa. In particular, these vector-valued Rademacher sums played an indispensable rôle in helping us arrive at the groups $G^{(\ell)}$ specified in §3, especially for $\ell > 3$, and also allowed us to formulate and test hypotheses regarding the modularity of the (vector-valued) functions $H_g^{(\ell)}$, including eta product expressions and the occurrences of classical mock theta functions; considerations which ultimately developed into the discussion of §4. A detailed discussion of the Rademacher construction is beyond the scope of this article but a full treatment is to be the focus of forthcoming work. The fact that the Rademacher sum approach proved so powerful may be taken as strong evidence in support of the *umbral moonshine conjecture*, Conjecture 5.4.

2 Automorphic Forms

In this section we discuss the modular objects that play a rôle in the connection between mock modular forms and finite groups that we will develop later in the paper. We also establish our notation and describe various relationships between Jacobi forms, theta functions, and vector-valued mock modular forms, including mock theta functions.

In what follows we take τ in the upper half-plane $\mathbb{H}$ and $z \in \mathbb{C}$, and adopt the shorthand notation $e(x) = e^{2\pi i x}$. We also define $q = e(\tau)$ and $y = e(z)$ and write

$$\gamma\tau = \frac{a\tau + b}{c\tau + d}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{2.1}$$

for the natural action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$ and

$$\gamma(\tau, z) = \left(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d} \right) \tag{2.2}$$

for the action of $SL_2(\mathbb{Z})$ on $\mathbb{H} \times \mathbb{C}$.

2.1 Mock Modular Forms

Mock theta functions were first introduced in 1920 by S. Ramanujan in his last letter to Hardy. This letter contained 17 examples divided into four of *order* 3, ten of *order* 5 and three of *order* 7. Ramanujan did not define what he meant by the term order and to this day there seems to be no universally agreed upon definition. In this paper we use the term order only as a historical label. Ramanujan wrote his mock theta functions as what he termed "Eulerian series" that today would be recognised as specialisations of q-hypergeometric series. A well studied example is the order 3 mock theta function

$$f(q) = 1 + \sum_{n=1}^{\infty} \frac{q^{n^2}}{(1+q)^2(1+q^2)^2 \cdots (1+q^n)^2} = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q;q)_n^2} \ , \tag{2.3}$$

where we have introduced the *q-Pochhammer* symbol

$$(a;q)_n = \prod_{k=0}^{n-1} (1 - aq^k) \,. \tag{2.4}$$

Interest in and applications of mock theta functions has burgeoned during the last decade following the work of Zwegers [44] who found an intrinsic definition of mock theta functions and their near modular behavior, and many applications of his work can be found in combinatorics [46, 47], characters of infinite-dimensional Lie superalgebras [48, 49], topological field theory [50, 51, 52, 53], the computation of quantum invariants of three-dimensional manifolds [54], and the counting of black hole states in string theory [45]. Descriptions of this breakthrough and some of the history of mock theta functions can be found in [55], [56] and [57].

Mock theta function are now understood as a special case of more general objects known as mock modular forms. A holomorphic function $h(\tau)$ on $\mathbb{H}$ is called a *(weakly holomorphic) mock* *modular form* of weight k for a discrete group Γ (e.g. a congruence subgroup of $SL_2(\mathbb{Z})$) if it has at most exponential growth as $\tau \to \alpha$ for any $\alpha \in \mathbb{Q}$, and if there exists a holomorphic modular form $f(\tau)$ of weight $2-k$ on Γ such that the *completion* of h given by

$$\hat{h}(\tau) = h(\tau) + (4i)^{k-1} \int_{-\bar{\tau}}^{\infty} (z+\tau)^{-k} \overline{f(-\bar{z})} \mathrm{d}z, \tag{2.5}$$

is a (non-holomorphic) modular form of weight k for Γ for some multiplier system ν say. In this case the function f is called the *shadow* of the mock modular form h. Even though h is not a modular form, it is common practice to call ν the multiplier system of h. One can show that ν is the conjugate of the multiplier system of f. In most of the examples in this paper we will deal with vector-valued mock modular forms so that the completion in fact transforms as $\nu(\gamma)\hat{h}(\gamma\tau)(c\tau+d)^{-k} = \hat{h}(\tau)$ in case the weight is k for all $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma$ where ν is a matrix-valued function on Γ.

The completion $\hat{h}(\tau)$ satisfies interesting differential equations. For instance, completions of mock modular forms were identified as weak Maass forms (non-holomorphic modular forms which are eigenfunctions of the Laplace operator) in [46] as a part of their solution to the longstanding Andrews–Dragonette conjecture. Note that we have the identity

$$2^{1-k}\pi\Im(\tau)^k \frac{\partial \hat{h}(\tau)}{\partial \bar{\tau}} = -2\pi i \overline{f(\tau)} \tag{2.6}$$

when f is the shadow of h.

Thanks to Zweger's work we may define a *mock theta function* to be a q-series $h = \sum_n a_n q^n$ such that for some $\lambda \in \mathbb{Q}$ the assignment $\tau \mapsto q^\lambda h|_{q=e(\tau)}$ defines a mock modular form of weight $1/2$ whose shadow is a unary (i.e. attached to a quadratic form in one variable) theta series of weight $3/2$.

In this paper we add one more rôle for mock theta functions to the list mentioned earlier; namely we conjecture that specific sets of mock theta functions appear as McKay–Thompson series associated to (also conjectural) infinite-dimensional modules for a sequence of groups $G^{(\ell)}$ which we refer to as the *umbral groups* and label by the *lambent* integers $\ell \in \Lambda = \{2,3,4,5,7,13\}$, which are just those positive integers that are one greater than a divisor of 12.

Many of the mock theta functions that appear later in this paper appear either in Ramanujan's last letter to Hardy or in his lost notebook [58]. These include an order 2 mock theta function

$$\mu(q) = \sum_{n\geq 0} \frac{(-1)^n\, q^{n^2} (q;q^2)_n}{(-q^2;q^2)_n^2} \tag{2.7}$$

and an order 8 mock theta function

$$U_0(q) = \sum_{n\geq 0} \frac{q^{n^2}(-q;q^2)_n}{(-q^4;q^4)_n}, \tag{2.8}$$

both of which appear at lambency 2 in connection with $G^{(2)} \simeq M_{24}$. The function $f(q)$ of (2.3) together with

$$\begin{aligned}
\phi(q) &= 1 + \sum_{n=1}^{\infty} \frac{q^{n^2}}{(1+q^2)(1+q^4)\cdots(1+q^{2n})} \\
\chi(q) &= 1 + \sum_{n=1}^{\infty} \frac{q^{n^2}}{(1-q+q^2)(1-q^2+q^4)\cdots(1-q^n+q^{2n})} \\
\omega(q) &= \sum_{n=0}^{\infty} \frac{q^{2n(n+1)}}{(1-q)^2(1-q^3)^2\cdots(1-q^{2n+1})^2} \\
\rho(q) &= \sum_{n=0}^{\infty} \frac{q^{2n(n+1)}}{(1+q+q^2)(1+q^3+q^6)\cdots(1+q^{2n+1}+q^{4n+2})}
\end{aligned} \tag{2.9}$$

constitute five order 3 mock theta functions appearing at lambency 3, and the four order 10 mock theta functions

$$\begin{aligned}
\phi_{10}(q) &= \sum_{n=0}^{\infty} \frac{q^{n(n+1)/2}}{(q;q^2)_{n+1}} \\
\psi_{10}(q) &= \sum_{n=0}^{\infty} \frac{q^{(n+1)(n+2)/2}}{(q;q^2)_{n+1}} \\
X(q) &= \sum_{n=0}^{\infty} \frac{(-1)^n q^{n^2}}{(-q;q)_{2n}} \\
\chi_{10}(q) &= \sum_{n=0}^{\infty} \frac{(-1)^n q^{(n+1)^2}}{(-q;q)_{2n+1}}
\end{aligned} \tag{2.10}$$

appear at lambency 5.

More mock theta functions were found later by others. At lambency 4 we will encounter order 8 mock theta functions discussed in [59] with q-expansions

$$\begin{aligned}
S_0(\tau) &= \sum_{n\geq 0} \frac{q^{n^2}(-q;q^2)_n}{(-q^2;q^2)_n}\\
S_1(\tau) &= \sum_{n\geq 0} \frac{q^{n(n+2)}(-q;q^2)_n}{(-q^2;q^2)_n}\\
T_0(\tau) &= \sum_{n\geq 0} \frac{q^{(n+1)(n+2)}(-q^2;q^2)_n}{(-q;q^2)_n}\\
T_1(\tau) &= \sum_{n\geq 0} \frac{q^{n(n+1)}(-q^2;q^2)_n}{(-q;q^2)_n}.
\end{aligned} \tag{2.11}$$

It is curious to note that the order is divisible by the lambency in each example.

2.2 Jacobi Forms

We now discuss Jacobi forms following [60]. We say a holomorphic function $\phi : \mathbb{H}\times\mathbb{C}\to\mathbb{C}$ is an *unrestricted Jacobi form* of weight k and index m for $SL_2(\mathbb{Z})$ if it transforms under the Jacobi group $SL_2(\mathbb{Z})\ltimes\mathbb{Z}^2$ as

$$\phi(\tau,z) = (c\tau+d)^{-k} e(-m\tfrac{cz^2}{c\tau+d})\,\phi(\gamma(\tau,z)) \tag{2.12}$$

$$\phi(\tau,z) = e(m(\lambda^2\tau+2\lambda z))\,\phi(\tau,z+\lambda\tau+\mu) \tag{2.13}$$

where $\gamma\in SL_2(\mathbb{Z})$ and $\lambda,\mu\in\mathbb{Z}$. In what follows we refer to the transformations (2.12) and (2.13) as the *modular* and *elliptic* transformations, respectively. The invariance of $\phi(\tau,z)$ under $\tau\to\tau+1$ and $z\to z+1$ implies a Fourier expansion

$$\phi(\tau,z) = \sum_{n,r\in\mathbb{Z}} c(n,r)q^n y^r \tag{2.14}$$

and the elliptic transformation can be used to show that $c(n,r)$ depends only on the *discriminant* r^2-4mn and r mod $2m$, and so we have $c(n,r)=C(r^2-4mn,\tilde{r})$ for some function $D\mapsto C(D,\tilde{r})$ where $\tilde{r}\in\{-m,\dots,m-1\}$, for example. An unrestricted Jacobi form is called a *weak Jacobi form*, a *(strong) Jacobi form*, or a *Jacobi cusp form* according as the Fourier coefficients satisfy $c(n,r)=0$ whenever $n<0$, $C(D,\tilde{r})=0$ whenever $D>0$, or $C(D,\tilde{r})=0$ whenever $D\geq 0$, respectively. In a slight departure from [60] we denote the space of weak Jacobi forms of weight k and index m by $J_{k,m}$.

In what follows we will need two further generalisations of the above definitions. The first is straightforward and replaces $SL_2(\mathbb{Z})$ by a finite index subgroup $\Gamma \subset SL_2(\mathbb{Z})$ in the modular transformation law. The second is more subtle and leads to *meromorphic Jacobi forms* which obey the modular and elliptic transformation laws but are such that the functions $z \mapsto \phi(\tau, z)$ are allowed to have poles lying at values of $z \in \mathbb{C}$ that map to torsion points of the elliptic curve $\mathbb{C}/(\mathbb{Z}\tau + \mathbb{Z})$. Our treatment of meromorphic Jacobi forms follows [44] and [45].

A property of (weak) Jacobi forms that will be important for us later is that they admit an expansion in terms of the *index m theta functions*,

$$\theta_r^{(m)}(\tau, z) = \sum_{n\in\mathbb{Z}} q^{(2mn+r)^2/4m} y^{2mn+r}, \tag{2.15}$$

given by

$$\phi(\tau, z) = \sum_{r (\mathrm{mod}\ 2m)} \tilde{h}_r(\tau)\theta_r^{(m)}(\tau, z) \tag{2.16}$$

in case the index of ϕ is m, where the *theta-coefficients* $\tilde{h}_r(\tau)$ constitute the components of a vector-valued modular form of weight $k - 1/2$ when k is the weight of ϕ. Recall that a vector valued function $\tilde{h} = (\tilde{h}_r)$ is called a *vector-valued modular form* of weight k for $\Gamma \subset SL_2(\mathbb{Z})$ if

$$\tilde{h}_r(\tau) = \frac{1}{(c\tau + d)^k} \sum_s \nu_{rs}(\gamma)\tilde{h}_s(\tau) \tag{2.17}$$

for all $\gamma \in \Gamma$ and $\tau \in \mathbb{H}$ for some matrix-valued function $\nu = (\nu_{rs})$ on Γ called the *multiplier system* for $\tilde{h}$.

The modular transformation law (2.12) with $\gamma = -I_2$ implies that $\phi(\tau, -z) = (-1)^k\phi(\tau, z)$. Combining this with the identity $\theta_{-r}^{(m)}(\tau, z) = \theta_r^{(m)}(\tau, -z)$ we see that $\tilde{h}_r(\tau) = (-1)^k\tilde{h}_{-r}(\tau)$ and in particular we can recover a weak Jacobi form of weight k and index m from the $m - 1$ theta-coefficients $\{\tilde{h}_1, \cdots, \tilde{h}_{m-1}\}$ in case k is odd.

In what follows we will encounter only weight 0 and weight 1 Jacobi forms; a typical such form will be denoted by $\phi(\tau, z)$ or $\psi(\tau, z)$ according as the weight is 0 or 1 and we will write the theta-expansion of a weight 1 Jacobi form ψ as

$$\psi(\tau, z) = \sum_{r=1}^{m-1} \tilde{h}_r(\tau)\hat{\theta}_r^{(m)}(\tau, z) \tag{2.18}$$

where $\hat{\theta}_r^{(m)}(\tau, z) = \theta_{-r}^{(m)}(\tau, z) - \theta_r^{(m)}(\tau, z)$ (cf. (2.15)) for $r \in \{1, 2, \cdots, m - 1\}$.

2.3 Meromorphic Jacobi Forms

We now explain a connection between the vector-valued mock modular forms we shall consider in this paper and meromorphic Jacobi forms. We specialize our discussion to weight 1 meromorphic Jacobi forms of a particular form that arise in our pairing of Jacobi forms with groups $G^{(\ell)}$, and later in our computation of McKay–Thompson series; namely, we consider meromorphic weight 1 and index m Jacobi forms which can be written as

$$\psi(\tau,z)=\Psi_{1,1}(\tau,z)\phi(\tau,z) \tag{2.19}$$

for some weight 0 index $m-1$ (holomorphic) weak Jacobi form ϕ where $\Psi_{1,1}$ is the specific meromorphic Jacobi form of weight 1 and index 1 given by

$$\Psi_{1,1}(\tau,z)=-i\,\frac{\theta_1(\tau,2z)\,\eta(\tau)^3}{(\theta_1(\tau,z))^2}=\frac{y+1}{y-1}-(y^2-y^{-2})q+\cdots . \tag{2.20}$$

We note that $\psi(\tau,z)$ has a simple pole at $z=0$ with residue $\phi(\tau,0)/\pi i$. Since $\phi(\tau,z)$ is a weak Jacobi form of weight 0 the function $\tau\mapsto\phi(\tau,0)$ is a modular form of weight 0 (with no poles at any cusps) and is hence equal to a constant; we denote this constant by

$$\chi=\phi(\tau,0). \tag{2.21}$$

It was shown by Zwegers [44] that meromorphic Jacobi forms have a modified theta-expansion (cf. (2.16)) in terms of vector-valued mock modular forms; in [45] this expansion was recast as follows. Define the averaging operator

$$\mathrm{Av}^{(m)}\Big[F(y)\Big]=\sum_{k\in\mathbb{Z}}q^{mk^2}y^{2mk}F(q^ky) \tag{2.22}$$

which takes a function of $y=e(z)$ with polynomial growth and returns a function of z which transforms like an index m Jacobi form under the elliptic transformations (2.13). Now define the *polar part* of $\psi=\Psi_{1,1}\phi$ by

$$\psi^P(\tau,z)=\chi\mathrm{Av}^{(m)}\Big[\frac{y+1}{y-1}\Big] \tag{2.23}$$

where $\chi=\phi(\tau,0)$ and define the *finite part* of ψ by

$$\psi^F(\tau,z)=\psi(\tau,z)-\psi^P(\tau,z). \tag{2.24}$$

The term finite is appropriate because with the polar part subtracted the finite part no longer has a pole at $z = 0$.

It follows from the analysis in [45] that $\psi^F(\tau, z)$ is a weight 1 index m *mock Jacobi form*, meaning that it has a theta-expansion

$$\psi^F(\tau, z) = \sum_{r=1}^{m-1} h_r(\tau)\hat{\theta}_r^{(m)}(\tau, z) \tag{2.25}$$

(cf. (2.18)) where the theta-coefficients h_r comprise the components of a vector-valued mock modular form of weight 1/2.

In the above we have again used the fact that ψ, ψ^P and hence also ψ^F pick up a minus sign under the transformation $z \to -z$. Moreover, the vector-valued mock modular forms obtained in this way always have shadow function given by the *unary theta series*

$$S_r^{(m)}(\tau) = \frac{1}{2\pi i}\frac{\partial}{\partial z}\theta_r^{(m)}(\tau, z)\bigg|_{z=0} = \sum_{n\in\mathbb{Z}}(2mn + r)q^{(2mn+r)^2/4m}. \tag{2.26}$$

To see why this is so, and for later use, we introduce the functions

$$\mu_j^{(m)}(\tau, z) = (-1)^{1+2j}\sum_{k\in\mathbb{Z}} q^{mk^2}y^{2mk}\frac{(yq^k)^{-2j} + (yq^k)^{-2j+1} + \cdots + (yq^k)^{1+2j}}{1 - yq^k} \tag{2.27}$$

for m a positive integer and $2j \in \{0, 1, \ldots, m-1\}$. Note that $\mu_0^{(m)}(\tau, z)$ is proportional to the polar part of ψ, as identified in (2.23),

$$\mu_0^{(m)}(\tau, z) = \mathrm{Av}^{(m)}\Big[\frac{y+1}{y-1}\Big]. \tag{2.28}$$

Remark 2.1. The function $\mu_0^{(2)}$ is closely related to the Appell-Lerch sum $\mu(t, z)$ which features prominently in [44].

The $\mu_0^{(m)}$ enjoy the following relation to the modular group $SL_2(\mathbb{Z})$. Define the *completion* of $\mu_0^{(m)}(\tau, z)$ by setting

$$\hat{\mu}^{(m)}(\tau, \bar{\tau}, z) = \mu_0^{(m)}(\tau, z) + \frac{1}{\sqrt{2m}}\frac{1}{(4i)^{1/2}}\sum_{r=-m}^{m-1}\theta_r^{(m)}(\tau, z)\int_{-\bar{\tau}}^{i\infty}(z + \tau)^{-1/2}\overline{S_r^{(m)}(-\bar{z})}\,\mathrm{d}z. \tag{2.29}$$

Then $\hat{\mu}^{(m)}$ transforms like a Jacobi form of weight 1 and index m for $SL_2(\mathbb{Z})$ but is not holomorphic. Therefore, from the transformation of the polar part

$$\psi^P(\tau,z)=\chi\mu_0^{(m)}(\tau,z) \tag{2.30}$$

we see that the shadow of $h=(h_r)$ is given by $\chi S^{(m)}=\big(\chi S_r^{(m)}\big)$, as we claimed. This means that the vector-valued mock modular forms arising in this way are closely related to mock theta functions: By the definition given in §2.1 we have $h_r(\tau)=q^{\lambda}M_r$ for some $\lambda\in\mathbb{Q}$ with M_r a mock theta function.

2.4 Superconformal Algebra

In §2.3 we saw how to associate a vector-valued mock modular form (h_r) to a weight 1 meromorphic Jacobi form ψ satisfying $\psi=\Psi_{1,1}\phi$ for some weak Jacobi form ϕ via the theta-expansion of the finite part of ψ. It will develop that the weight 0 forms ϕ of relevance to us have a close relation to the representation theory of the 2-dimensional $N=4$ superconformal algebra. To see this recall (cf. [16, 17, 18]) that this algebra contains subalgebras isomorphic to the affine Lie algebra $\hat{\mathfrak{sl}}_2$ and the Virasoro algebra, and in a unitary representation the former of these acts with level $m-1$, for some integer $m>1$, and the latter with central charge $c=6(m-1)$, and the unitary irreducible highest weight representations $V_{h,j}^{(m)}$ are labelled by the two *quantum numbers* h and j which are the eigenvalues of L_0 and $\frac{1}{2}J_0^3$, respectively, when acting on the highest weight state. (We adopt a normalisation of the $SU(2)$ current J^3 such that the zero mode J_0^3 has integer eigenvalues.) In the Ramond sector of the superconformal algebra there are two types of highest weight representations: the *massless* (or *BPS*, *supersymmetric*) ones with $h=\frac{m-1}{4}$ and $j\in\{0,\frac{1}{2},\cdots,\frac{m-1}{2}\}$, and the *massive* (or *non-BPS*, *non-supersymmetric*) ones with $h>\frac{m-1}{4}$ and $j\in\{\frac{1}{2},1,\cdots,\frac{m-1}{2}\}$. Their *(Ramond) characters*, defined as

$$\mathrm{ch}_{h,j}^{(m)}(\tau,z)=\mathrm{tr}_{V_{h,j}^{(m)}}\left((-1)^{J_0^3}y^{J_0^3}q^{L_0-c/24}\right), \tag{2.31}$$

are given by

$$\mathrm{ch}^{(m)}_{\frac{m-1}{4},j}(\tau,z)=(\Psi_{1,1}(\tau,z))^{-1}\mu_j^{(m)}(\tau,z) \tag{2.32}$$

and

$$\mathrm{ch}_{h,j}^{(m)}(\tau,z)=(-1)^{2j+1}(\Psi_{1,1}(\tau,z))^{-1}\,q^{h-\frac{m-1}{4}-\frac{j^2}{m}}\,\hat\theta_{2j}^{(m)}(\tau,z) \tag{2.33}$$

in the massless and massive cases, respectively, [17] where the function $\mu_j^{(m)}(\tau,z)$ is defined as in (2.27) and $\hat\theta_r^{(m)}(\tau,z)$ is as in the sentence following (2.18).

We can use the above results to derive a decomposition of an arbitrary weight 0 index $m-1$

Jacobi form $\phi(\tau,z)$ into $N=4$ characters as follows. Set $\psi(\tau,z)=\Psi_{1,1}(\tau,z)\phi(\tau,z)$ as in §2.3 and write

$$h_r(\tau)=\sum_n c_r(n-r^2/4m)q^{n-r^2/4m} \tag{2.34}$$

for the Fourier expansion of the theta-coefficient h_r of the finite part ψ^F of ψ (cf. (2.25)). Then use (2.24), (2.25) and (2.30) along with (2.32) and (2.33) to obtain

$$\begin{aligned}\phi =&\chi\,\mathrm{ch}^{(m)}_{\frac{m-1}{4},0}+\sum_{r=1}^{m-1}(-1)^{r+1}c_r(-\tfrac{r^2}{4m})\left(\mathrm{ch}^{(m)}_{\frac{m-1}{4},\frac{r}{2}}+2\mathrm{ch}^{(m)}_{\frac{m-1}{4},\frac{r-1}{2}}+\mathrm{ch}^{(m)}_{\frac{m-1}{4},\frac{r-2}{2}}\right)\\ &+\sum_{r=1}^{m-1}(-1)^{r+1}\sum_{n=1}^{\infty}c_r(n-\tfrac{r^2}{4m})\,\mathrm{ch}^{(m)}_{\frac{m-1}{4}+n,\frac{r}{2}}\end{aligned} \tag{2.35}$$

where $\chi=\phi(\tau,0)$ is the constant such that $\chi S^{(m)}$ is the shadow of $h=(h_r)$ (cf. §2.3). In deriving (2.35) we have used the relation

$$\mu^{(m)}_{\frac{r-2}{2}}+2\mu^{(m)}_{\frac{r-1}{2}}+\mu^{(m)}_{\frac{r}{2}}=(-1)^{r+1}\,q^{-\frac{r^2}{4m}}\hat{\theta}^{(m)}_r \tag{2.36}$$

subject to the understanding that $\mu^{(m)}_{-\frac{1}{2}}=\mathrm{ch}^{(m)}_{\frac{m-1}{4},-\frac{1}{2}}=0$.

One way a weak Jacobi form of weight 0 and index $m-1$ having integer coefficients can arise in nature is as the elliptic genus of a 2-dimensional $N=4$ superconformal field theory with central charge $c=6(m-1)$. There is a unique weak Jacobi form of weight 0 and index 1 up to scale and this (suitably scaled) turns out to be the elliptic genus of a superconformal field theory attached to a $K3$ surface. Then the above analysis at $m=2$ recovers the mock modular form $H^{(2)}$ (the vectors have $m-1=1$ components in this case) exhibiting the connection to the Mathieu group M_{24} observed by Eguchi–Ooguri–Tachikawa in [15].

In the next section we will construct a distinguished family of extremal weight 0 weak Jacobi forms $\phi^{(\ell)}(\tau,z)$ with corresponding vector-valued weight 1/2 mock modular forms $H^{(\ell)}(\tau)$ according to the procedure (2.19-2.25) of §2.3. In §3 we will specify finite groups $G^{(\ell)}$ for which the forms $H^{(\ell)}$ will serve as generating functions for the graded dimensions of conjectural bi-graded infinite-dimensional $G^{(\ell)}$-modules.

2.5 Extremal Jacobi Forms

In this section we identify the significance of the divisors of 12 from the point of view of the $N=4$ superconformal algebra. We introduce the notion of an extremal (weak) Jacobi form of weight 0 and integral index and we prove that the space $J^{\mathrm{ext}}_{0,m-1}$ of extremal forms with index $m-1$ has dimension 1 if $m-1$ divides 12, and has dimension 0 otherwise. Several of the extremal Jacobi forms we identify have appeared earlier in the literature in the context of studying decompositions of elliptic genera of Calabi–Yau manifolds [43, 61, 48] and in a recent study of the connection between black hole counting in superstring theory and mock modular forms [45].

For m a positive integer and ϕ a weak Jacobi form with weight 0 and index $m-1$ say ϕ is *extremal* if it admits an expression

$$\phi = a_{\frac{m-1}{4},0}\mathrm{ch}^{(m)}_{\frac{m-1}{4},0} + a_{\frac{m-1}{4},\frac{1}{2}}\mathrm{ch}^{(m)}_{\frac{m-1}{4},\frac{1}{2}} + \sum_{0<r<m}\sum_{\substack{n\in\mathbb{Z}\\ r^2-4mn<0}} a_{\frac{m-1}{4}+n,\frac{r}{2}}\mathrm{ch}^{(m)}_{\frac{m-1}{4}+n,\frac{r}{2}} \tag{2.37}$$

for some $a_{h,j}\in\mathbb{C}$ (cf. (2.35)) where the $N=4$ characters are as defined in (2.32) and (2.33). Write $J^{\mathrm{ext}}_{0,m-1}$ for the subspace of $J_{0,m-1}$ consisting of extremal weak Jacobi forms. Note that the extremal condition restricts both the massless and massive $N=4$ representations that can appear, for generally there are non-zero massive characters $\mathrm{ch}^{(m)}_{\frac{m-1}{4}+n,\frac{r}{2}}$ with $r^2-4mn\geq 0$.

We observe here that the extremal condition has a very natural interpretation in terms of the mock modular forms of weight 1/2 attached to weak Jacobi forms of weight 0 via the procedure detailed in §2.3. By comparing with (2.35) we find that the condition (2.37) on a weak Jacobi form ϕ of index $m-1$ is equivalent to requiring that the corresponding vector-valued mock modular form (h_r) obtained from the theta-expansion (2.25) of the finite part of the weight 1 Jacobi form $\psi=\Psi_{1,1}\phi$ has a single polar term $q^{-\frac{1}{4m}}$ in the first component h_1 and has all other components vanishing as $\tau\to i\infty$.

Our main result in this section is the following characterisation of extremal Jacobi forms.

Theorem 2.2. *If m is a positive integer then* $\dim J^{\mathrm{ext}}_{0,m-1}=1$ *in case* $m-1$ *divides* 12 *and* $\dim J^{\mathrm{ext}}_{0,m-1}=0$ *otherwise.*

In preparation for the proof of Theorem 2.2 we now summarise (aspects of) a useful construction given in [43]. The graded ring

$$J_{0,*}=\bigoplus_{m\geq 1} J_{0,m-1} \tag{2.38}$$

of (weak) Jacobi forms of weight 0 and integral index has an ideal

$$J_{0,*}(q) = \bigoplus_{m>1} J_{0,m-1}(q) = \left\{ \phi \in J_{0,*} \mid \phi(\tau,z) = \sum_{\substack{n,r\in\mathbb{Z}\\ n>0}} c(n,r) q^n y^r \right\} \tag{2.39}$$

consisting of Jacobi forms that vanish in the limit as $\tau \to i\infty$ (i.e. have vanishing coefficient of $q^0 y^r$, for all r, in their Fourier development). This ideal is principal and generated by a weak Jacobi form of weight 0 and index 6 given by

$$\zeta(\tau,z) = \frac{\theta_1(\tau,z)^{12}}{\eta(\tau)^{12}} \tag{2.40}$$

(cf. §A for θ_1 and η). Gritsenko shows [43] that for any positive integer m the quotient $J_{0,m-1}/J_{0,m-1}(q)$ is a vector space of dimension $m-1$ admitting a basis consisting of weight 0 index $m-1$ weak Jacobi forms $\varphi_n^{(m)}$ (denoted $\psi_{0,m-1}^{(n)}$ in [43]) for $1 \leq n \leq m-1$ such that the coefficient of $q^0 y^k$ in $\varphi_n^{(m)}$ vanishes for $|k| > n$ but does not vanish for $|k| = n$. In fact Gritsenko works in the subring $J_{0,*}^{\mathbb{Z}}$ of Jacobi forms having integer Fourier coefficients and his $\varphi_n^{(m)}$ furnish a $\mathbb{Z}$-basis for the $\mathbb{Z}$-module $J_{0,*}^{\mathbb{Z}}/J_{0,*}^{\mathbb{Z}}(q)$.

We show now that there are no non-zero extremal Jacobi forms in the ideal $J_{0,*}(q)$.

Lemma 2.3. *If ϕ is an extremal weak Jacobi form belonging to $J_{0,*}(q)$ then $\phi = 0$.*

Proof. If ϕ belongs to $J_{0,*}(q)$ then the coefficients of $q^0 y^k$ in the Fourier development of ϕ vanish for all k. This implies the vanishing of $a_{\frac{m-1}{4},0}$ and $a_{\frac{m-1}{4},\frac{1}{2}}$ in (2.37) where $m-1$ is the index of ϕ, and this in turn implies that the meromorphic Jacobi form $\psi = \Psi_{1,1}\phi$ of weight 1 and index m coincides with its finite part $\psi = \psi^F = \sum_r h_r \hat{\theta}_r^{(m)}$ (cf. (2.21), (2.23), (2.25)) and has theta-coefficients h_r that remain bounded as $\tau \to i\infty$. In particular, ψ is a (strong) Jacobi form of weight 1 and integral index but the space of such forms vanishes according to [62], so ϕ must vanish also. ☐

As a consequence of Lemma 2.3 we obtain that there are no extremal Jacobi forms with vanishing massless contribution at spin $j = 1/2$.

Lemma 2.4. *If ϕ is an extremal weak Jacobi form of index $m-1$ such that $a_{\frac{m-1}{4},\frac{1}{2}} = 0$ in (2.37) then $\phi = 0$.*

Proof. If ϕ is of weight 0 and index $m-1$ satisfying (2.37) then $\phi = a + b(y + y^{-1}) + O(q)$ as

$\tau \to i\infty$ where $a = a_{\frac{m-1}{4},0}$ and $b = -a_{\frac{m-1}{4},\frac{1}{2}}$. Following [43] we set

$$\tilde{\phi}(\tau, z) = \exp(-8\pi^2(m-1)G_2(\tau)z^2)\phi(\tau, z) \tag{2.41}$$

where $G_2(\tau) = -\frac{1}{24} + \sum_{n>0} \sigma_1(n)q^n$ is the unique up to scale mock modular form of weight 2 for $SL_2(\mathbb{Z})$ (with shadow a constant function) and consider the *Taylor expansion* $\tilde{\phi}(\tau, z) = \sum_{n\geq 0} f_n(\tau)z^n$. Then $f_0(\tau) = \chi = a + 2b$ and more generally $f_n(\tau)$ is a modular form of weight n for $SL_2(\mathbb{Z})$. In particular $f_2(\tau) = 0$. On the other hand the constant term of $f_2(\tau)$ is $\frac{1}{3}\pi^2(m-1)\chi - 4\pi^2 b$ so $(m-1)\chi = 12b$. By hypothesis $b = 0$ so $\chi = 0$ and ϕ belongs to $J_{0,m-1}(q)$ and thus vanishes according to Lemma 2.3. □

Applying Lemma 2.4 we obtain that the dimension of $J^{\mathrm{ext}}_{0,m-1}$ is at most 1 for any m.

Proposition 2.5. *We have* $\dim J^{\mathrm{ext}}_{0,m-1} \leq 1$ *for any positive integer* m.

Proof. If ϕ' and ϕ'' are two elements of $J^{\mathrm{ext}}_{0,m-1}$ then there exists $c \in \mathbb{C}$ such that $\phi = \phi' - c\phi''$ is extremal and has vanishing $a_{\frac{m-1}{4},\frac{1}{2}}$ in (2.37). Then $\phi = 0$ according to Lemma 2.4 and thus ϕ' belongs to the linear span of ϕ''. □

We now present a result which ultimately shows that there are only finitely many m for which $J^{\mathrm{ext}}_{0,m-1} \neq \{0\}$, and also illustrates the depth of the characterisation problem. In preparation for it let us write $S_2(m)$ for the space of cusp forms of weight 2 for $\Gamma_0(m)$ and recall that $f \in S_2(m)$ is called a *newform* if it is a Hecke eigenform, satisfying $T(n)f = \lambda_f(n)f$ whenever $(n, m) = 1$ for some $\lambda_f(n) \in \mathbb{C}$, that is uniquely specified (up to scale) in $S_2(m)$ by its Hecke eigenvalues $\{\lambda_f(n) \mid (n, m) = 1\}$ (cf. [63, §IX.7]). Write $S_2^{\mathrm{new}}(m)$ for the subspace of $S_2(m)$ spanned by newforms. Given $f \in S_2(m)$ define $L(f, s) = \sum_{n>0} a_f(n)n^{-s}$ for $\Re(s) > 1$ when $f(\tau) = \sum_{n>0} a_f(n)q^n$. Then $L(f, s)$ is called the *Dirichlet L-function* attached to f and admits an analytic continuation to $s \in \mathbb{C}$ (cf. [63, §IX.4] or [64, §3.6].) To a newform f in $S_2(m)$ (i.e. a Hecke eigenform in $S_2^{\mathrm{new}}(m)$) Eichler–Shimura theory attaches an Elliptic curve E_f defined over $\mathbb{Q}$ (cf. [63, §XI.11]), and the Birch–Swinnerton-Dyer conjecture predicts that E_f has rational points of infinite order whenever $L(f, s)$ vanishes at $s = 1$ (cf. [65, 66]).

Theorem 2.6. *If* m *is a positive integer and* $J^{\mathrm{ext}}_{0,m-1} \neq \{0\}$ *then* $L(f, 1) = 0$ *for every* $f \in S_2^{\mathrm{new}}(m)$.

Proof. Let m be a positive integer. For $L = \sqrt{2m}\mathbb{Z}$ the spaces $H^+_{k,L}$ and S_{2-k,L^-} of [67, §3] are naturally isomorphic to $\mathbb{J}_{k+1/2,m}$ and $S^{\mathrm{skew}}_{5/2-k,m}$, respectively, where $\mathbb{J}_{k,m}$ denotes the space of weak (pure) mock Jacobi forms of weight k and index m (cf. [45, §7.2]) and $S^{\mathrm{skew}}_{k,m}$ denotes the space of cuspidal skew-holomorphic Jacobi forms of weight k and index m (cf. [68]). Translating the construction (3.9) of [67] into this language, and taking $k = 1/2$, we obtain a pairing

$$\{\cdot\,,\cdot\} : S^{\mathrm{skew}}_{2,m} \times \mathbb{J}_{1,m} \to \mathbb{C}. \tag{2.42}$$

If $J^{\mathrm{ext}}_{0,m-1} \neq \{0\}$ then let $\phi \in J^{\mathrm{ext}}_{0,m-1}$ be a non-zero extremal Jacobi form and set $\psi = \Psi_{1,1}\phi$ as in §2.3. Then the finite part ψ^F of ψ belongs to $\mathbb{J}_{1,m}$ and the shadow of ψ^F (cf. [45, §7.2]) is the cuspidal skew-holomorphic Jacobi form $\chi\sigma^{(m)} \in S^{\mathrm{skew}}_{2,m}$ where $\chi = \phi(\tau, 0)$ and

$$\sigma^{(m)}(\tau, z) = \sum_{r \bmod 2m} \overline{S^{(m)}_r(\tau)}\theta^{(m)}_r(\tau, z). \tag{2.43}$$

Consider the linear functional λ_ϕ on $S^{\mathrm{skew}}_{2,m}$ defined by setting $\lambda_\phi(\varphi) = \{\varphi, \psi^F\}$. According to the definition of (2.42) (i.e. (3.9) of [67]) we have

$$\lambda_\phi(\varphi) = C_m\langle\varphi, \chi\sigma^{(m)}\rangle \tag{2.44}$$

for some (non-zero) constant C_m (depending only on m) where $\langle\cdot\,,\cdot\rangle$ denotes the Petersson inner product on $S^{\mathrm{skew}}_{2,m}$ (cf. [68]). On the other hand, inspection reveals that $h_{\pm1}(\tau) = \mp a_{\frac{m-1}{4},\frac{1}{2}}q^{-1/4m}(1 + O(q))$ when $\psi^F = \sum_{r \bmod 2m} h_r\theta^{(m)}_r$ so

$$\lambda_\phi(\varphi) = -2a_{\frac{m-1}{4},\frac{1}{2}}C_\varphi(1, 1) \tag{2.45}$$

according to Proposition 3.5 of [67] where $\varphi(\tau, z) = \sum C_\varphi(\Delta, r)\bar{q}^{\Delta/4m}q^{r^2/4m}y^r$ is the Fourier expansion of φ (cf. [68]). Applying Lemma 2.4 we deduce from (2.45) that $\lambda_\phi(\varphi) = 0$ for $\varphi \in S^{\mathrm{skew}}_{2,m}$ if and only if $C_\varphi(1, 1) = 0$.

Now let f be a newform in $S_2(m)$. If $\Lambda(f, s) = (2\pi)^{-s}m^{s/2}\Gamma(s)L(f, s)$ then either $\Lambda(f, s) = \Lambda(f, 2-s)$ or $\Lambda(f, s) = -\Lambda(f, 2-s)$. In the latter case $L(f, 1)$ necessarily vanishes. In the former case Theorem 1 of [69] attaches a skew-holomorphic Jacobi form $\varphi_f \in P^{\mathrm{skew}}_{2,m}$ to f having the same eigenvalues as f under the Hecke operators $T(n)$ for $(n, m) = 1$. (See [69] for the action of Hecke operators on skew-holmorphic Jacobi forms.) Here $P^{\mathrm{skew}}_{2,m}$ denotes the subspace of $S^{\mathrm{skew}}_{2,m}$ spanned by cuspidal Hecke eigenforms which are not of the trivial type: say $\varphi \in S^{\mathrm{skew}}_{2,m}$ is of the *trivial type* if the Fourier coefficients $C_\varphi(\Delta, r)$ are non-vanishing only when Δ is a perfect square. Cusp forms of the trivial type are orthogonal to $P^{\mathrm{skew}}_{2,m}$ with respect to the Petersson inner product (cf. the proof of Theorem 3 in [69]) and $\sigma^{(m)}$ is a cusp form of the trivial type, so we have $\lambda_\phi(\varphi_f) = 0$ by virtue of (2.44) and hence $C_{\varphi_f}(1, 1) = 0$ according to the previous paragraph. The proposition now follows by taking $(\Delta, r) = (1, 1)$ in the formula at the bottom of p. 503 of [68]. ☐

Applying Theorem 2.6 we see, for example, that there are no non-zero extremal Jacobi forms of index 10, for it is known that $L(f, 1) \neq 0$ for $f = \eta(\tau)^2\eta(11\tau)^2 \in S_2(11)$. (Cf. [68, p. 504] where the corresponding Jacobi form φ_f is computed explicitly.)

Duke showed [70] that there is a constant C such that at least $Cm/\log^2 m$ newforms $f \in S_2(m)$ satisfy $L(f, 1) \neq 0$ when m is a sufficiently large prime, and Ellenberg proved an effective version of this result [42] that is valid also for composite m. We will apply the formulas of Ellenberg momentarily in order to deduce an upper bound on the m for which $J^{\rm ext}_{0,m-1} \neq \{0\}$, but since these formulas give better results for values of m with fewer divisors we first show that $J^{\rm ext}_{0,m-1}$ must vanish whenever m is divisible by more than one prime, and whenever $m = p^\nu$ is a prime power with $\nu > 2$.

Proposition 2.7. *If m is divisible by more than one prime then $J^{\rm ext}_{0,m-1} = \{0\}$.*

Proof. Suppose that $m = p^\nu d$ with p prime, $\nu > 0$, $d > 1$ and $(p, d) = 1$. With notation as in the proof of Theorem 2.6 we suppose $\phi \in J^{\rm ext}_{0,m-1}$ is non-zero and consider the linear functional $\lambda_\phi : S^{\rm skew}_{2,m} \to \mathbb{C}$. The Atkin–Lehner involution W_{p^∞} (cf. [45, §4.4]) acts on $S^{\rm skew}_{2,m}$ in such a way that

$$(W_{p^\infty}\varphi)(\tau, z) = \sum_{r \bmod 2m} \overline{g_{r^*}(\tau)}\theta^{(m)}_r(\tau, z) \tag{2.46}$$

when $\varphi = \sum_{r \bmod 2m} \overline{g_r}\theta^{(m)}_r$ is the theta-decomposition of $\varphi \in S^{\rm skew}_{2,m}$ and where the map $r \mapsto r^*$ is defined so that r^* is the unique solution (mod $2m$) to $r^* \equiv -r \pmod{2p^\nu}$ and $r^* \equiv r \pmod{2d}$.

By the definition of λ_ϕ we have $\lambda_\phi(\varphi) = \{\varphi, \psi^F\} = \overline{\chi}\sum_r\langle g_r, S^{(m)}_r\rangle$ with φ as above since the shadow of ψ^F is $\chi\sigma^{(m)} = \chi\sum_r \overline{S^{(m)}_r}\theta^{(m)}_r$. Taking $\varphi = W_{p^\infty}\sigma^{(m)}$ we see that $\lambda_\phi(W_{p^\infty}\sigma^{(m)}) = \overline{\chi}\sum_r\langle S^{(m)}_{r^*}, S^{(m)}_r\rangle$ is not zero according to the computation

$$\langle S^{(m)}_{r^*}, S^{(m)}_r\rangle = \frac{1}{2^6\pi m^{3/2}}\left(\delta_{r^*,r \pmod{2m}} - \delta_{r^*,-r \pmod{2m}}\right) \tag{2.47}$$

which can be obtained using the Rankin–Selberg formula. (Cf., e.g., (3.15) of [45]. A very similar computation is carried out in Proposition 10.2 of [45].) On the other hand for $\varphi = W_{p^\infty}\sigma^{(m)}$ we have $C_\varphi(1, 1) = 0$ since $r^* \not\equiv \pm1 \pmod{2m}$ for $r \equiv \pm1 \pmod{2m}$, and $S^{(m)}_r = O(q^{2/4m})$ for $r \not\equiv \pm1 \pmod{2m}$. So $\lambda_\phi(W_{p^\infty}\sigma^{(m)}) = 0$ in light of (2.45). This contradiction proves the claim. ☐

Proposition 2.8. *If m is a prime power $m = p^\nu$ and $\nu > 2$ then $J^{\rm ext}_{0,m-1} = \{0\}$.*

Proof. The proof is very similar to that of Proposition 2.7, computing $\lambda_\phi(\varphi)$ two ways using (2.44) and (2.45), but taking now $\varphi(\tau,z) = \sigma^{(p)}(\tau,p^\mu z)$ or $\varphi(\tau,z) = \sigma^{(p^2)}(\tau,p^\mu z)$ according as $\nu = 2\mu+1$ or $\nu = 2\mu+2$ with $\mu > 0$. The transformation $\varphi(\tau,z) \mapsto \varphi(\tau,tz)$ maps Jacobi forms of index m to Jacobi forms of index t^2m, and is one of the Hecke-like operators of [45, §4.4]. We leave the remaining details—the expression of φ in terms of $S^{(m)}_r$ and $\theta^{(m)}_r$, and an application of (2.47)—to the reader. ☐

We require one more result in advance of the proof of Theorem 2.2.

Lemma 2.9. *If $m = p^2$ for some prime p and there exists a newform $f \in S_2(p)$ with $L(f,1) \neq 0$ then $J^{\rm ext}_{0,m-1} = \{0\}$.*

Proof. Let m, p and $f \in S_2(p)$ be as in the statement of the lemma. Then, as in the proof of Theorem 2.6, there exists a uniquely determined $\varphi_f \in P^{\rm skew}_{2,p}$ with the same Hecke eigenvalues as f according to the results of [68, 69], and since $L(f,1) \neq 0$ we have $C_{\varphi_f}(1,1) \neq 0$. Now set $\varphi'_f = \varphi_f|V_p$ where V_p is the Hecke-like operator (cf. [45, (4.37)]) that maps Jacobi forms of index m to forms of index pm, and whose action is given explicitly by

$$C_{\varphi|V_p}(\Delta,r) = \sum_{d|\left(\frac{\Delta-r^2}{4p^2},r,p\right)} dC_\varphi\left(\frac{\Delta}{d^2},\frac{r}{d}\right) \tag{2.48}$$

in our special case that $\varphi = \varphi_f \in J^{\rm skew}_{2,p}$ when $\varphi = \sum C_\varphi(\Delta,r)\bar{q}^{\Delta/4p}q^{r^2/4p}y^r$ is the Fourier expansion of φ.

To prove that $J^{\rm ext}_{0,m-1} = \{0\}$ suppose that ϕ is a non-zero extremal form of index $m-1 = p^2-1$ and consider the functional $\lambda_\phi : S^{\rm skew}_{2,p^2} \to \mathbb{C}$ constructed in the proof of Theorem 2.6. Since $C_{\varphi_f}(1,1) \neq 0$ we have $C_{\varphi'_f}(1,1) \neq 0$ according to (2.48). So $\lambda_\phi(\varphi'_f) \neq 0$ thanks to (2.45) and Lemma 2.4. On the other hand the Hecke-like operator V_p commutes with all Hecke operators $T(n)$ (with, in our case, $(n,p)=1$) and so restricts to a map $P^{\rm skew}_{2,p} \to P^{\rm skew}_{2,p^2}$. Thus φ'_f is a Hecke-eigenform in $P^{\rm skew}_{2,p^2}$ and we have $\lambda_\phi(\varphi'_f) = 0$ by virtue of (2.44) and the fact that $P^{\rm skew}_{2,p^2}$ is orthogonal to the cusp form $\sigma^{(p^2)}$ of the trivial type. This contradiction completes the proof. ☐

Proof of Theorem 2.2. We have that $\dim J^{\rm ext}_{0,m-1} \leq 1$ for all m according to Proposition 2.5, and $\dim J^{\rm ext}_{0,m-1} = 0$ unless $m = p$ or $m = p^2$ for some prime p by virtue of Propositions 2.7 and 2.8. Applying Theorem 1 of [42] with $\sigma = 13/25\pi$, for example, shows that some $f \in S^{\rm new}_2(p)$ satisfies $L(f,1) \neq 0$ whenever p is a prime greater than or equal to 3001. The tables of arithmetic data on newforms posted at [71] give the non-vanishing or otherwise for every newform of level up to 5134, and inspecting the entries for the 430 prime levels less than 3000 we see that $S_2(p)$ has a newform f with $L(f,1)\neq 0$ for every prime p so long as $S_2(p)$ has a newform. We have $S_2^{\rm new}(p) = S_2(p) = \{0\}$ just when $p\in\{2,3,5,7,13\}$, so we conclude from Theorem 2.6 that $J^{\rm ext}_{0,m-1}=\{0\}$ for every prime $m=p$ such that $p-1$ does not divide 12. Also, we conclude from Lemma 2.9 that $J^{\rm ext}_{0,m-1}=\{0\}$ whenever $m=p^2$ is the square of a prime and $p-1$ does not divide 12. We inspect the tables of [71] again to find newforms with non-vanishing critical central value at levels $7^2=49$ and $13^2=169$, and apply Theorem 2.6 to conclude $J^{\rm ext}_{0,48}=J^{\rm ext}_{0,168}=\{0\}$.

So we now require to determine the dimension of $J^{\rm ext}_{0,m-1}$ for $m\in\{2,3,4,5,7,9,13,25\}$, which are the primes and squares of primes m for which $S_2(m)=\{0\}$. For this we utilise the basis $\{\varphi_n^{(m)}\}$ for $J_{0,*}$ determined by Gritsenko [43], and discussed, briefly, after the statement of Theorem 2.2 above. In more detail, the ring $J_{0,*}$ is finitely generated, by $\varphi_1^{(2)}$, $\varphi_1^{(3)}$ and $\varphi_1^{(4)}$, where

$$\begin{aligned}\varphi_1^{(2)} &= 4\left(f_2^2+f_3^2+f_4^2\right),\\ \varphi_1^{(3)} &= 2\left(f_2^2f_3^2+f_3^2f_4^2+f_4^2f_2^2\right),\\ \varphi_1^{(4)} &= 4f_2^2f_3^2f_4^2,\end{aligned}\tag{2.49}$$

and $f_i(\tau,z)=\theta_i(\tau,z)/\theta_i(\tau,0)$ for $i\in\{2,3,4\}$ (cf. §A.2). If we work over $\mathbb{Z}$ then we must include $\varphi_1^{(5)}$

$$\varphi_1^{(5)}=\frac{1}{4}\left(\varphi_1^{(4)}\varphi_1^{(2)}-(\varphi_1^{(3)})^2\right)\tag{2.50}$$

as a generator also, so that $J^{\mathbb{Z}}_{0,*}=\mathbb{Z}[\varphi_1^{(2)},\varphi_1^{(3)},\varphi_1^{(4)},\varphi_1^{(5)}]$. Following [43] we define

$$\begin{aligned}\varphi_1^{(7)} &= \varphi_1^{(3)}\varphi_1^{(5)}-(\varphi_1^{(4)})^2,\\ \varphi_1^{(9)} &= \varphi_1^{(3)}\varphi_1^{(7)}-(\varphi_1^{(5)})^2,\\ \varphi_1^{(13)} &= \varphi_1^{(5)}\varphi_1^{(9)}-2(\varphi_1^{(7)})^2,\end{aligned}\tag{2.51}$$

and define $\varphi_1^{(m)}$ for the remaining positive integers m according to the following recursive procedure. For $(12,m-1)=1$ and $m>5$ we set

$$\varphi_1^{(m)}=(12,m-5)\varphi_1^{(m-4)}\varphi_1^{(5)}+(12,m-3)\varphi_1^{(m-2)}\varphi_1^{(3)}-2(12,m-4)\varphi_1^{(m-3)}\varphi_1^{(4)}.\tag{2.52}$$

For $(12,m-1)=2$ and $m>10$ we set

$$\varphi_1^{(m)}=\frac{1}{2}\big((12,m-5)\varphi_1^{(m-4)}\varphi_1^{(5)}+(12,m-3)\varphi_1^{(m-2)}\varphi_1^{(3)}-2(12,m-4)\varphi_1^{(m-3)}\varphi_1^{(4)}\big).\tag{2.53}$$

For $(12, m-1) = 3$ and $m > 9$ we set

$$\varphi_1^{(m)} = \frac{2}{3}(12, m-4)\varphi_1^{(m-3)}\varphi_1^{(4)} + \frac{1}{3}(12, m-7)\varphi_1^{(m-6)}\varphi_1^{(7)} - (12, m-5)\varphi_1^{(m-4)}\varphi_1^{(5)}. \tag{2.54}$$

For $(12, m-1) = 4$ and $m > 16$ we set

$$\varphi_1^{(m)} = \frac{1}{4}\big((12, m-13)\varphi_1^{(m-12)}\varphi_1^{(13)} + (12, m-5)\varphi_1^{(m-4)}\varphi_1^{(5)} - (12, m-9)\varphi_1^{(m-8)}\varphi_1^{(9)}\big). \tag{2.55}$$

For $(12, m-1) = 6$ and $m > 18$ we set

$$\varphi_1^{(m)} = \frac{1}{3}(12, m-4)\varphi_1^{(m-3)}\varphi_1^{(4)} + \frac{1}{6}(12, m-7)\varphi_1^{(m-6)}\varphi_1^{(7)} - \frac{1}{2}(12, m-5)\varphi_1^{(m-4)}\varphi_1^{(5)}. \tag{2.56}$$

Finally, for $(12, m-1) = 12$ and $m > 24$ we set[1]

$$\varphi_1^{(m)} = \frac{1}{6}(12, m-4)\varphi_1^{(m-3)}\varphi_1^{(4)} - \frac{1}{4}(12, m-5)\varphi_1^{(m-4)}\varphi_1^{(5)} + \frac{1}{12}(12, m-7)\varphi_1^{(m-6)}\varphi_1^{(7)}. \tag{2.57}$$

The $\varphi_2^{(m)}$ are defined by setting

$$\begin{aligned}
\varphi_2^{(3)} &= (\varphi_1^{(2)})^2 - 24\,\varphi_1^{(3)},\\
\varphi_2^{(4)} &= \varphi_1^{(2)}\varphi_1^{(3)} - 18\,\varphi_1^{(4)},\\
\varphi_2^{(5)} &= \varphi_1^{(2)}\varphi_1^{(4)} - 16\,\varphi_1^{(5)},
\end{aligned} \tag{2.58}$$

and

$$\varphi_2^{(m)} = (12, m-4)\,\varphi_1^{(m-3)}\varphi_1^{(4)} - (12, m-5)\varphi_1^{(m-4)}\varphi_1^{(5)} - (12, m-1)\varphi_1^{(m)} \tag{2.59}$$

for $m > 5$, and the remaining $\varphi_n^{(m)}$ for $2 \leq m \leq 25$ are given by

$$\begin{aligned}
\varphi_n^{(m)} &= \varphi_{n-1}^{(m-3)}\varphi_1^{(4)},\\
\varphi_{m-2}^{(m)} &= (\varphi_1^{(2)})^{m-3}\varphi_1^{(3)},\\
\varphi_{m-1}^{(m)} &= (\varphi_1^{(2)})^{m-1},
\end{aligned} \tag{2.60}$$

where the first equation of (2.60) holds for $3 \leq n \leq m-3$.

Next we calculate that $\varphi_1^{(m)}$ defines a non-zero element of $J_{0,m-1}^{\rm ext}$ when $m-1$ divides 12, so $\dim J_{0,m-1}^{\rm ext} = 1$ for $m \in \{2, 3, 4, 5, 7, 13\}$, and it remains to show that $J_{0,m-1}^{\rm ext} = \{0\}$ for $m \in \{9, 25\}$. To do this we first observe that a weight zero form with vanishing Fourier coefficients of q^0y^k for $|k|>1$ necessarily has the form

---

[1] Our expression gives one possible correction for a small error in the last line on p. 10 of [43].

$$c_{0,1}\,\varphi_1^{(m)}+\sum_{i=1}^{\lfloor\frac{m-1}{6}\rfloor}\sum_{j=1}^{m-6i-1}c_{i,j}\,\zeta^i\varphi_j^{(m-6i)} \tag{2.61}$$

for some $c_{i,j}\in\mathbb{C}$. One has to show that the only such linear combination satisfying the extremal condition (2.37) necessarily has $c_{i,j}=0$ for all i,j (including $(i,j)=(0,1)$). An explicit inspection of the coefficients of terms q^ny^r with $r^2-4mn\geq 0$ for $n=1,2,3,4$ is sufficient to establish that $\dim J^{\rm ext}_{0,m-1}=0$ when $m=9$. For $m=25$ we notice that there are at least as many possible polar terms in the above sum as the number of parameters $c_{i,j}$. Indeed, by explicitly solving the system of linear equations we find that there is no accidental cancellation and the only solution to (2.61) that also satisfies the extremal condition (2.37) is zero. In this way we conclude that there is no solution to (2.37) in $J_{0,m-1}$ for $m\in\{9,25\}$. This completes the proof of the Theorem. □

The quantity $Z^{(2)}(\tau,z)=2\varphi_1^{(2)}(\tau,z)$ is equal to the elliptic genus of a(ny) $K3$ surface and the factor of two relating $Z^{(2)}$ to $\varphi_1^{(2)}$ is required in the $K3/M_{24}$ connection in order for the mock modular form $H^{(2)}=(H_1^{(2)})$ derived from $Z^{(2)}$ to have coefficients compatible with an interpretation as dimensions of representations of M_{24} for which the corresponding McKay–Thompson series $H_g^{(2)}$ have integer Fourier coefficients. Inspired by this we define the *umbral Jacobi forms*

$$Z^{(\ell)}(\tau,z)=2\varphi_1^{(\ell)}(\tau,z) \tag{2.62}$$

for $\ell\in\Lambda=\{2,3,4,5,7,13\}$. We also set $\chi^{(\ell)}$ to be the constant $Z^{(\ell)}(\tau,0)$ and find that $\chi^{(\ell)}=24/(\ell-1)$ for all $\ell\in\Lambda$.

Remark 2.10. The identity $(m-1)\chi=12b$ for $\phi=a+b(y+y^{-1})+O(q)$ established in Lemma 2.4 shows that there is no such Jacobi form with $b=1$ and a an integer unless $m-1$ divides 12. In particular, there is no extremal Jacobi form $\phi\in J^{\rm ext}_{0,m-1}$ such that $a_{\frac{m-1}{4},\frac12}=1$ and $a_{\frac{m-1}{4},0}$ is an integer unless $m-1$ divides 12. Inspecting the $Z^{(\ell)}$ we conclude that the divisors of 12 are exactly the values of $m-1$ for which such a Jacobi form exists.

Following the discussion in §§2.3,2.4, each of the umbral Jacobi forms $Z^{(\ell)}$ leads to an $(\ell-1)$-vector-valued mock modular form $H^{(\ell)}=\left(H_r^{(\ell)}\right)$ through the decomposition of $Z^{(\ell)}$ into $N=4$ characters, or equivalently through the decomposition of $\psi^{(\ell)}=\Psi_{1,1}Z^{(\ell)}$ into its polar and finite parts and the theta-expansion of the finite part,

$$\psi^{(\ell),F}(\tau,z)=\sum_{r=1}^{\ell-1}H_r^{(\ell)}(\tau)\hat{\theta}_r^{(\ell)}(\tau,z). \tag{2.63}$$

As we have explained above, the extremal condition translates into a natural condition on the vector-valued mock modular forms $H^{(\ell)}$: The requirement that $Z^{(\ell)}$ takes the form (2.37) is equivalent to requiring that the only polar term in $H^{(\ell)}=\left(H_r^{(\ell)}\right)$ is $-2q^{-\frac{1}{4\ell}}$ in the component $H_1^{(\ell)}$ and all the other components $H_r^{(\ell)}$ for $r\neq 1$ vanish as $\tau\to i\infty$, so that

$$H_r^{(\ell)}(\tau)=-2\delta_{r,1}q^{-\frac{1}{4\ell}}+O(q^{\frac{1}{4\ell}}) \tag{2.64}$$

for $0<r<\ell$.

Some low order terms in the Fourier expansions of the component functions $H_r^{(\ell)}$ obtained from the extremal forms $Z^{(\ell)}$ for $\ell\in\{2,3,4\}$ are given as follows.

$$H_1^{(2)}(\tau)=2q^{-1/8}\left(-1+45q+231q^2+770q^3+2277q^4+5796q^5+\cdots\right) \tag{2.65}$$

$$\begin{aligned}H_1^{(3)}(\tau)&=2q^{-1/12}\left(-1+16q+55q^2+144q^3+330q^4+704q^5+\cdots\right)\\H_2^{(3)}(\tau)&=2q^{2/3}\left(10+44q+110q^2+280q^3+572q^4+1200q^5+\cdots\right)\end{aligned} \tag{2.66}$$

$$\begin{aligned}H_1^{(4)}(\tau)&=2q^{-1/16}\left(-1+7q+21q^2+43q^3+94q^4+168q^5+\cdots\right)\\H_2^{(4)}(\tau)&=2q^{3/4}\left(8+24q+56q^2+112q^3+216q^4+392q^5+\cdots\right)\\H_3^{(4)}(\tau)&=2q^{7/16}\left(3+14q+28q^2+69q^3+119q^4+239q^5+\cdots\right)\end{aligned} \tag{2.67}$$

As a prelude to the next section we note that the coefficients 16, 55 and 144 appearing in $H_1^{(3)}$ are dimensions of irreducible representations of the Mathieu group M_{12} and that $\chi^{(3)}=12$ is the dimension of the defining permutation representation of M_{12} just as $\chi^{(2)}=24$ is the dimension of the defining permutation representation of M_{24}. We further note that the low-lying coefficients appearing in $H_2^{(3)}$ are dimensions of faithful irreducible representations of a group $2.M_{12}$. Here the notation $2.G$ denotes a group with a $\mathbb{Z}/2\mathbb{Z}$ normal subgroup such that $2.G/(\mathbb{Z}/2\mathbb{Z})=G$. Note also that $2.M_{12}$ has a faithful (and irreducible) 12-dimensional signed permutation representation (cf. §3.2). The pattern that the coefficients of $H_r^{(\ell)}$ for r odd are dimensions of representations of a group $\bar{G}^{(\ell)}$ with a permutation representation of dimension $\chi^{(\ell)}$ and the coefficients of $H_r^{(\ell)}$ for r even are dimensions of faithful representations of a group $G^{(\ell)}=2.\bar{G}^{(\ell)}$ with a signed permutation representation of degree $\chi^{(\ell)}$ persists for all $\ell\in\Lambda$; detailed descriptions of the (unsigned) permutation and signed permutation representations of

$\bar{G}^{(\ell)}$ and $G^{(\ell)}$ are given in §3.3 and §3.4.

The leading terms in the q-expansions of the mock modular forms $H_1^{(\ell)}$ that correspond via the procedure described earlier to the Jacobi forms $Z^{(\ell)}$ for $\ell \in \{5, 7, 13\}$ are

$$\begin{aligned} H_1^{(5)}(\tau) &= 2q^{-1/20}\left(-1+4q+9q^2+20q^3+35q^4+60q^5+\cdots\right),\\ H_1^{(7)}(\tau) &= 2q^{-1/28}\left(-1+2q+3q^2+5q^3+10q^4+15q^5+21q^6+\cdots\right),\\ H_1^{(13)}(\tau) &= 2q^{-1/52}\left(-1+q+q^2+q^4+q^5+2q^6+3q^7+\cdots\right). \end{aligned} \tag{2.68}$$

To avoid clutter we refrain from describing low order terms in the q-expansions of $H_r^{(\ell)}$ for $\ell \in \{5, 7, 13\}$ and $r > 1$ here, but these can be read off from the 1A entries in Tables 27-29, 31-35, and 37-47.

We conclude this section with a comparison of our condition (2.37) with other notions of extremal in the literature. A notion of *extremal holomorphic conformal field theory* was given in [27] following earlier related work on vertex operator superalgebras in [72]. Extremal CFT's have central charge $c = 24k$ with k a positive integer and are defined in such a way that their partition function satisfies

$$Z = q^{-k}\prod_{n>1}^{\infty}\frac{1}{1-q^n} + O(q)\prod_{n>0}^{\infty}\frac{1}{1-q^n} \tag{2.69}$$

as $\tau \to i\infty$ where the term $O(q)$ is presumed to be a series in q with integer coefficients, so that the second term $O(q)\prod_{n>0}(1-q^n)^{-1}$ in (2.69) represents an integer combination of irreducible characters of the Virasoro algebra. The first term in (2.69) is the vacuum character of the Virasoro algebra, generated by the vacuum state, and thus any Virasoro primaries above the vacuum are only allowed to contribute positive powers of q to the partition function. At this time the only known example of an extremal CFT is that determined by the monster vertex operator algebra $V^\natural$ whose partition function has $k = 1$ in (2.69) so this notion is, at the very least, good at singling out extraordinary structure.

If ϕ is a weak Jacobi form of weight 0 and index $m-1$ that is extremal in our sense then we have

$$\phi = a_{\frac{m-1}{4},0}\mathrm{ch}^{(m)}_{\frac{m-1}{4},0} + a_{\frac{m-1}{4},\frac{1}{2}}\mathrm{ch}^{(m)}_{\frac{m-1}{4},\frac{1}{2}} + \sum_{0<r<m} O(q)\frac{\hat{\theta}_r^{(m)}}{\Psi_{1,1}} \tag{2.70}$$

as $\tau \to i\infty$ for some $a_{\frac{m-1}{4},0}$ and $a_{\frac{m-1}{4},\frac{1}{2}}$, and the third term in (2.70) is a natural counterpart to the second term in (2.69) since the massive $N = 4$ characters (in the Ramond sector with $(-1)^F$ insertion) are all of the form $q^\alpha\hat{\theta}_r^{(m)}/\Psi_{1,1}$ for some α and some r. It is harder to argue that the massless $N = 4$ contributions in (2.70) are in direct analogy with the first term in

(2.69) since the vacuum state in a superconformal field theory with $N = 4$ supersymmetry will generally (i.e. when the index is greater than 1) give rise to non-zero massless $N = 4$ character contributions with spin greater than 1/2. We remark that the stronger condition

$$\phi = a_{\frac{m-1}{4},0}\mathrm{ch}^{(m)}_{\frac{m-1}{4},0} + \sum_{0<r<m} O(q)\frac{\hat{\theta}^{(m)}_r}{\Psi_{1,1}} \tag{2.71}$$

has no solutions according to Lemma 2.4. According to Theorem 2.2 the six Jacobi forms $Z^{(\ell)}$ of umbral moonshine are the unique solutions to (2.70) having $a_{\frac{m-1}{4},\frac{1}{2}} = -2$.

A notion of *extremal conformal field theory with* $N = (2,2)$ *superconformal symmetry* was introduced in [73] and the Ramond sector partition function of such an object defines a weak Jacobi form of weight 0 and some index but will generally not coincide with a Jacobi form that is extremal in our sense since, as has been mentioned, our functions are free from contributions arising from states in the Ramond sector that are related by spectral flow to the vacuum state in the Neveu–Schwarz sector, except in the case of index 1.

Despite the absence of a notion of extremal conformal field theory underlying our extremal Jacobi forms it is interesting to reflect on the fact that the one known example of an extremal CFT is that (at $k = 1$) determined by the moonshine vertex operator algebra of [12] with partition function $Z(\tau) = J(\tau)$. This function encodes dimensions of irreducible representations of the monster and, as we have seen to some extent above and will see in more detail below, the quantities which play the same role with respect to the groups $G^{(\ell)}$ of umbral moonshine are precisely the mock modular forms $H^{(\ell)}(\tau) = \left(H^{(\ell)}_r(\tau)\right)$ which arise naturally from the extremal Jacobi forms $Z^{(\ell)}(\tau, z)$ according to the procedure of §2.3. One of the main motivations for the notion of extremal CFT introduced in [27] was a possible connection to pure (chiral) gravity theory in 3-dimensional Anti-de Sitter space via the AdS/CFT correspondence and it is interesting to compare this with the important rôle that Rademacher sum constructions play in monstrous moonshine [25], umbral moonshine at $\ell = 2$ [24], and in umbral moonshine more generally (cf. §1 and §5.2). We refer to [28, 29, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85] for further discussions on the AdS/CFT correspondence and extremal CFT's.

2.6 Siegel Modular Forms

Siegel modular forms are automorphic forms which generalise modular forms by replacing the modular group $SL_2(\mathbb{Z})$ by the *genus n Siegel modular group* $Sp_{2n}(\mathbb{Z})$ and the upper half-plane $\mathbb{H}$

by the *genus n Siegel upper half-plane* $\mathbb{H}_n$. Here we restrict ourselves to the case $n = 2$. Define

$$J = \begin{pmatrix} 0 & -I_2 \\ I_2 & 0 \end{pmatrix} \tag{2.72}$$

with I_2 the unit 2×2 matrix and let $Sp_4(\mathbb{Z})$ be the group of 4×4 matrices $\gamma \in M_4(\mathbb{Z})$ satisfying $\gamma J \gamma^t = J$. If we write γ in terms of 2×2 matrices with integer entries A, B, C, D,

$$\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{2.73}$$

then the condition $\gamma J \gamma^t = J$ becomes

$$AB^t = BA^t, \qquad CD^t = DC^t, \qquad AD^t - BC^t = 1. \tag{2.74}$$

Just as $SL_2(\mathbb{Z})$ has a natural action on $\mathbb{H}$, the (genus 2) Siegel modular group $Sp_4(\mathbb{Z})$ has a natural action on the (genus 2) Siegel upper half-plane, $\mathbb{H}_2$, defined as the set of 2×2 complex, symmetric matrices

$$\Omega = \begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} \tag{2.75}$$

obeying

$$\mathrm{Im}(\tau) > 0, \qquad \mathrm{Im}(\sigma) > 0, \qquad \det(\mathrm{Im}(\Omega)) > 0. \tag{2.76}$$

The action of $\gamma \in Sp_4(\mathbb{Z})$ is given by

$$\gamma\Omega = (A\Omega + B)(C\Omega + D)^{-1} \tag{2.77}$$

when γ is given by (2.73). A *Siegel modular form* of weight k for $\Gamma \subset Sp_4(\mathbb{Z})$ is a holomorphic function $F : \mathbb{H}_2 \to \mathbb{C}$ obeying the transformation law

$$F((A\Omega + B)(C\Omega + D)^{-1}) = \det(C\Omega + D)^k F(\Omega) \tag{2.78}$$

for $\gamma \in \Gamma$. We can write a *Fourier-Jacobi decomposition* of $F(\Omega)$ in terms of $p = e(\sigma)$ as

$$F(\Omega) = \sum_{m=0}^{\infty} \varphi_m(\tau, z) p^m \tag{2.79}$$

and the transformation law for $F(\Omega)$ can be used to show that the Fourier-Jacobi coefficient $\varphi_m(\tau,z)$ is a Jacobi form of weight k and index m. (This is one of the main motivations for the notion of Jacobi form; cf. [86] for an early analysis with applications to affine Lie algebras.) We can thus write

$$\varphi_m(\tau,z)=\sum_{\substack{n,r\in\mathbb{Z}\\ 4mn-r^2\geq 0}} c(m,n,r)q^ny^r \tag{2.80}$$

and then

$$F(\Omega)=\sum_{m,n,r} c(m,n,r)p^mq^ny^r. \tag{2.81}$$

A special class of Siegel modular forms, called *Spezialschar* by Maass, arise by taking the Jacobi-Fourier coefficients φ_m to be given by $\varphi_m=\varphi_1|V_m$ where $\varphi_1\in J_{k,1}$ and V_m is the *Hecke-like operator* defined so that if $\varphi_1=\sum_{n,r}c(n,r)q^ny^r$ then

$$\varphi_1|V_m=\sum_{n,r}\left(\sum_{j|(n,r,m)} j^{k-1}c\left(\frac{nm}{j^2},\frac{r}{j}\right)\right)q^ny^r. \tag{2.82}$$

It develops that the function

$$F(\Omega)=\sum_{m=0}^{\infty}(\varphi_1|V_m)(\tau,z)p^m \tag{2.83}$$

is a weight k Siegel modular form known as the *Saito–Kurokawa* or *additive lift* of the weight k and index 1 Jacobi form φ_1. An important example arises by taking the additive lift of the Jacobi form

$$\varphi_{10,1}(\tau,z)=-\eta(\tau)^{18}\theta_1(\tau,z)^2 \tag{2.84}$$

(cf. §A) which produces the *Igusa cusp form*

$$\Phi_{10}(\Omega)=\sum_{m=1}^{\infty}(\varphi_{10,1}|V_m)(\tau,z)p^m \tag{2.85}$$

The Igusa cusp form also admits a product representation obtained from the *Borcherds* or *exponential lift* of the umbral Jacobi form $Z^{(2)}=2\varphi_{2,1}$ (cf. (2.49)) which is

$$\Phi_{10}(\Omega)=pqy\prod_{\substack{m,n,r\in\mathbb{Z}\\ (m,n,r)>0}}(1-p^mq^ny^r)^{c^{(2)}(4mn-r^2)} \tag{2.86}$$

where the coefficients $c^{(2)}$ are defined by the Fourier expansion of $Z^{(2)}$

$$Z^{(2)}(\tau,z)=\sum_{n,r}c^{(2)}(4n-r^2)q^ny^r \tag{2.87}$$

and where the condition $(m,n,r)>0$ is that either $m>0$ or $m=0$ and $n>0$ or $m=n=0$ and $r<0$.

According to Theorem 2.1 in [87] the umbral Jacobi form $Z^{(\ell)}$ defines a Siegel modular form $\Phi^{(\ell)}$ on the *paramodular group* $\Gamma^+_{\ell-1}<Sp_4(\mathbb{Q})$ via the Borcherds lift for each $\ell\in\Lambda$. (We refer the reader to [87] for details.) For small values of ℓ these Siegel forms $\Phi^{(\ell)}$ have appeared in the literature (in particular in the work [88, 87] of Gritsenko–Nikulin). Taking $\ell\in\Lambda$ and defining $c^{(\ell)}(n,r)$ so that

$$Z^{(\ell)}(\tau,z)=\sum_{n,r}c^{(\ell)}(n,r)q^ny^r \tag{2.88}$$

we have the exponential lift $\Phi^{(\ell)}$ of weight $k=c^{(\ell)}(0,0)/2$ for $\Gamma^+_{\ell-1}$ given by

$$\Phi^{(\ell)}(\Omega)=p^{A(\ell)}q^{B(\ell)}y^{C(\ell)}\prod_{(m,n,r)>0}(1-p^mq^ny^r)^{c^{(\ell)}(mn,r)} \tag{2.89}$$

where

$$A(\ell)=\frac{1}{24}\sum_{r}c^{(\ell)}(0,r),\quad B(\ell)=\frac{1}{2}\sum_{r>0}rc^{(\ell)}(0,r),\quad C(\ell)=\frac{1}{4}\sum_{r}r^2c^{(\ell)}(0,r). \tag{2.90}$$

Note that for $\ell\in\{2,3,4,5\}$ we have $\Phi^{(\ell)}=(\Delta_k)^2$ in the notation of [87] where k is given by $k=(7-\ell)/(\ell-1)$. These four functions Δ_k appear as denominator functions for Lorentzian Kac–Moody Lie (super)algebras in [87, §5.1] (see also [89]) and in connection with mirror symmetry for $K3$ surfaces in [87, §5.2]. We refer the reader to §5.5 for more discussion on this.

3 Finite Groups

In this section we introduce the *umbral groups* $G^{(\ell)}$ for $\ell\in\Lambda=\{2,3,4,5,7,13\}$. It will develop that the representation theory of $G^{(\ell)}$ is intimately related to the vector-valued mock modular form $H^{(\ell)}$ of §2.5.

We specify the abstract isomorphism types of the groups explicitly in §3.1. Each group admits a quotient $\bar{G}^{(\ell)}$ which is naturally realised as a permutation group on $24/(\ell-1)$ points. We construct these permutations explicitly in §3.3. In order to construct the $G^{(\ell)}$ we use signed permutations—a notion we discuss in §3.2—and explicit generators for the $G^{(\ell)}$ are finally obtained in §3.4. The signed permutation constructions are then used (again in §3.4) to define characters for $G^{(\ell)}$—the *twisted Euler characters*—which turn out to encode the automorphy of the mock modular forms $H_g^{(\ell)}$ that are be described in detail in §4. In §§3.5-3.6 we describe curious connections between the $G^{(\ell)}$ and certain Dynkin diagrams.

3.1 Specification

For $\ell \in \Lambda$ let abstract groups $G^{(\ell)}$ and $\bar{G}^{(\ell)}$ be as specified in Table 1. We will now explain the notation used therein (moving from right to left).

We write n as a shorthand for $\mathbb{Z}/n\mathbb{Z}$ (in the second and third rows of Table 1) and Sym_n denotes the symmetric group on n points. We write Alt_n for the alternating group on n points, which is the subgroup of Sym_n consisting of all even permutations.

We say that G is a *double cover* of a group H and write $G \simeq 2.H$ (cf. [90, §5.2]) in case G has a subgroup Z of order 2 that is normal (and therefore central, being of order two) with the property that G/Z is isomorphic to H. We say that G is a *non-trivial* double cover of H if it is a double cover that is not isomorphic to the direct product $2 \times H$. (We don't usually write $G \simeq 2.H$ unless G is a non-trivial cover of H.) The cyclic group of order 4 is a non-trivial double cover of the group of order 2.

For n a positive integer and q a prime power we write $GL_n(q)$ for the general linear group of invertible $n \times n$ matrices with coefficients in the finite field $\mathbb{F}_q$ with q elements, and we write $SL_n(q)$ for the subgroup consisting of matrices with determinant 1. We write $PGL_n(q)$ for the quotient of $GL_n(q)$ by its centre but we adopt the ATLAS convention (cf. [90, §2.1]) of writing $L_n(q)$ as a shorthand for $PSL_n(q)$—being the quotient of $SL_n(q)$ by its centre—since this group is typically simple. In the case that $q = 3$ and n is even the centre of $SL_n(q)$ has order 2 and $SL_n(q) \simeq 2.L_n(q)$ is a non-trivial double cover of $L_n(q)$.

In case $q = 5$ the centre of $GL_n(q)$ is a cyclic group of order 4 and so $GL_n(5)$ has a unique central subgroup of order 2. We write $GL_n(5)/2$ for the quotient of $GL_n(5)$ by this subgroup. In case $n = 2$ the exceptional isomorphism $PGL_2(5) \simeq Sym_5$ tells us then that $GL_2(5)/2$ is a double cover of the symmetric group Sym_5. For $n \geq 4$ there are two isomorphism classes of non-trivial double covers of Sym_n with the property that the central subgroup of order 2 is contained in the commutator subgroup of the double cover—these are the so-called *Schur double covers* (cf. [90, §4.1, §6.7]). The group $GL_2(5)/2$ is curious in that its only subgroup of index 2 is a direct product $2 \times Alt_5$ and its commutator subgroup is a copy of Alt_5, so it is a non-trivial double cover of Sym_5 that is not isomorphic to either of the Schur double covers.

The symbols $AGL_n(q)$ denote the *affine linear group* generated by the natural action of

Table 1: The Umbral Groups

ℓ	2	3	4	5	7	13
$G^{(\ell)}$	M_{24}	$2.M_{12}$	$2.AGL_3(2)$	$GL_2(5)/2$	$SL_2(3)$	4
$\bar{G}^{(\ell)}$	M_{24}	M_{12}	$AGL_3(2)$	$PGL_2(5)$	$L_2(3)$	2

$GL_n(q)$ on $\mathbb{F}_q^n$ and the translations $x \mapsto x + v$ for $v \in \mathbb{F}_q^n$. The group $AGL_n(q)$ may be realised as a subgroup of $GL_{n+1}(q)$ so in particular $AGL_3(2)$ embeds in $GL_4(2)$. The group $GL_4(2)$ is unusual amongst the $GL_n(2)$ in that it admits a non-trivial double cover $2.GL_4(2)$, which is a manifestation of the exceptional isomorphism $GL_4(2) \simeq Alt_8$. There are two conjugacy classes of subgroups of $2.Alt_8$ of order 2688, which is twice the order of $AGL_3(2)$. The groups in both classes are isomorphic to a particular non-trivial double cover of $AGL_3(2)$ and this is the group we denote $2.AGL_3(2)$ in Table 1.

The symbols M_{24} and M_{12} denote the sporadic simple Mathieu groups attached to the binary and ternary Golay codes, respectively. The group M_{24} is the automorphism group of the binary Golay code, which is the unique self-dual linear binary code of length 24 with minimum weight 8 (cf. [91]). The ternary Golay code is the unique self-dual linear ternary code of length 12 with minimum weight 6 (cf. [91]) and its automorphism group is a non-trivial double cover of M_{12}. There is a unique such group up to isomorphism (cf. [90]) which we denote by $2.M_{12}$ in Table 1.

Observe that for $\ell > 2$ the group $G^{(\ell)}$ has a unique central subgroup of order 2. Then $\bar{G}^{(\ell)}$ may be described for $\ell > 2$ by $\bar{G}^{(\ell)} = G^{(\ell)}/2$ and $G^{(\ell)}$ is a non-trivial double cover of $\bar{G}^{(\ell)}$ for each $\ell > 2$. It will develop in §3.2 that $\bar{G}^{(\ell)}$ is a permutation group of degree $n = 24/(\ell - 1)$ for each $\ell \in \Lambda$. Since $G^{(2)} \simeq M_{24}$ has trivial centre and is already a permutation group of degree $24 = 24/(2 - 1)$ we set $\bar{G}^{(2)} = G^{(2)}$.

Generalising the notation used above for double covers we write $G \simeq N.H$ to indicate that G fits into a short exact sequence

$$1 \to N \to G \to H \to 1 \tag{3.1}$$

for some groups N and H. We write $G \simeq N{:}H$ for a group of the form $N.H$ for which the sequence (3.1) splits (i.e. in case G is a semi-direct product $N \rtimes H$). Then we have $AGL_3(2) \simeq 2^3{:}L_2(7)$ according to the exceptional isomorphism $L_3(2) \simeq L_2(7)$ where 2^3 denotes an elementary abelian group of order 8.

3.2 Signed Permutations

The group of *signed permutations of degree* n or *hyperoctahedral group of degree* n, denoted here by Oct_n, is a semi-direct product $2^n{:}Sym_n$ where 2^n denotes an elementary abelian group of order 2^n. We may realise it explicitly as the subgroup of invertible linear transformations of an n-dimensional vector space generated by sign changes and permutations in a chosen basis. (This recovers $2^n{:}Sym_n$ so long as the vector space is defined over a field of characteristic other than 2. In case the characteristic is 2 the sign changes are trivial and we recover Sym_n.) When we write Oct_n we usually have in mind the data of fixed subgroups $N \simeq 2^n$ and $H \simeq Sym_n$ such that N is normal and $Oct_n/N \simeq H$, such as are given by sign changes and coordinate permutations, respectively, when we realise Oct_n explicitly as described above. In what follows we assume such data to be chosen for each n and write $Oct_n \to Sym_n$ for the composition $Oct_n \to Oct_n/N \simeq Sym_n$.

We will show in §3.4 that each $G^{(\ell)}$ for $\ell \in \Lambda$ admits a realisation as a subgroup of Oct_n for $n = 24/(\ell-1)$ such that the image of $G^{(\ell)}$ under the map $Oct_n \to Sym_n$ is $\bar{G}^{(\ell)}$.

In §3.4 we use the following modification of the usual cycle notation for permutations to denote elements of Oct_n. Suppose Ω is a set with n elements and V is the vector space generated over a field $\mathbf{k}$ by the symbols $\{e_x\}_{x\in\Omega}$. Whereas the juxtaposition xy occurring in a cycle $(...xy...)$ indicates a coordinate permutation mapping e_x to e_y, we write $x\bar{y}$ to indicate that a sign change is applied so that e_x is mapped to $-e_y$. Then, for example, if $\Omega = \{\infty, 0, 1, 2, 3, 4\}$ we write $\sigma = (\bar{\infty}2\bar{3}4)(\bar{0})$ for the element of $GL(V)$ determined by

$$\sigma : e_\infty \mapsto e_2, \quad e_2 \mapsto -e_3, \quad e_3 \mapsto e_4, \quad e_4 \mapsto -e_\infty, \quad e_0 \mapsto -e_0. \tag{3.2}$$

(We think of the bar over the 3 in $(...2\bar{3}4...)$ as "occurring between" the 2 and 3.) We call the symbol 1 a *fixed point* of σ and we call 0 an *anti-fixed point*.

We define the *signed permutation character* of Oct_n by setting

$$\begin{aligned} \chi : Oct_n &\to \mathbb{Z} \\ g &\mapsto \chi_g = h_{g,+} - h_{g,-} \end{aligned} \tag{3.3}$$

where $h_{g,+}$ is the number of fixed points of g and $h_{g,-}$ is the number of anti-fixed points. Then χ is just the character of the ordinary representation of Oct_n furnished by V that obtains when the field $\mathbf{k}$ is taken to be $\mathbb{C}$. Writing $\bar{\chi}$ for the composition of $Oct_n \to Sym_n$ with the usual permutation character of Sym_n we have

$$\begin{aligned} \bar{\chi} : Oct_n &\to \mathbb{Z} \\ g &\mapsto \bar{\chi}_g = h_{g,+} + h_{g,-} \end{aligned} \tag{3.4}$$

and we call $\bar{\chi}$ the *unsigned permutation character* of Oct_n. Then the signed and unsigned permutation characters χ and $\bar{\chi}$ together encode the number of fixed and anti-fixed points for each $g \in Oct_n$.

Let us take $\mathbf{k} = \mathbb{C}$ in the realisation of Oct_n as a subgroup of $GL(V)$. Then to each element $g \in Oct_n$ we assign a *signed permutation Frame shape* Π_g which encodes the eigenvalues of the corresponding (necessarily diagonalisable) linear transformation of V in the following way. An expression

$$\Pi_g = \prod_{k \geq 1} k^{m_g(k)} \tag{3.5}$$

with $m_g(k) \in \mathbb{Z}$ and $m_g(k) = 0$ for all but finitely many k indicates that g defines a linear transformation with $m_g(k)$ eigenvalues equal to $\mathrm{e}(j/k)$ for each $0 \leq j < k$ in the case that all the $m_g(k)$ are non-negative. If some $m_g(k)$ are negative then we find the number of eigenvalues equal to ξ say by looking at how many copies of ξ are present in $\Pi_g^+ = \prod_{k \geq 1,\, m_g(k) > 0} k^{m_g(k)}$ and subtracting the number of copies indicated by $\Pi_g^- = \prod_{k \geq 1,\, m_g(k) < 0} k^{-m_g(k)}$. Observe that the signed permutation character of Oct_n is recovered from the signed permutation Frame shapes via $\chi_g = m_g(1)$. Also, the Frame shape is invariant under conjugacy.

The map $Oct_n \to Sym_n$ (which may be realised by "ignoring" sign changes) furnishes a (non-faithful) permutation representation of Oct_n on n points; we call it the *unsigned permutation representation* of Oct_n. We write $g \mapsto \bar{\Pi}_g$ for the corresponding cycle shapes and call them the *unsigned permutation Frame shapes* attached to Oct_n. Observe that the unsigned permutation character of Oct_n is recovered from the unsigned permutation Frame shapes via $\bar{\chi}_g = \bar{m}_g(1)$ when $\bar{\Pi}_g = \prod_{k \geq 1} k^{\bar{m}_g(k)}$.

Observe that we obtain a faithful permutation representation $Oct_n \to Sym_{2n}$ by regarding $g \in Oct_n$ as a permutation of the $2n$ points $\{\pm e_x\}_{x \in \Omega}$. We denote the corresponding cycle shapes $\tilde{\Pi}_g$ and call them the *total permutation Frame shapes* attached to Oct_n. Then the total permutation Frame shapes can be recovered from the signed and unsigned permutation Frame shapes for if we define a formal product on Frame shapes by setting

$$\Pi\,\Pi' = \prod_{k\geq 1} k^{m(k)+m'(k)}. \tag{3.6}$$

in case $\Pi = \prod_{k\geq 1} k^{m(k)}$ and $\Pi' = \prod_{k\geq 1} k^{m'(k)}$ then we have $\tilde{\Pi}_g = \Pi_g \bar{\Pi}_g$ for all $g \in Oct_n$.

Suppose that G is a subgroup of Oct_n with the property that the intersection $G \cap N$ has order 2. (This will be the case for $G = G^{(\ell)}$ when $\ell \in \Lambda$ and $\ell \neq 2$.) Let z be the unique and necessarily central involution in $G \cap N$ and write $\bar{G}$ for the image of G under the map $Oct_n \to Sym_n$. Then the conjugacy class of zg depends only on the conjugacy class of g and we obtain an involutory map $[g] \mapsto [zg]$ on the conjugacy classes of G. We call $[g]$ and $[zg]$ *paired* conjugacy classes and we say that $[g]$ is *self-paired* if $[g] = [zg]$. Observe that z must act as -1 times the identity in any faithful irreducible representation of G so if $g \mapsto \chi(g)$ is the character of such a representation then $\chi(zg) = -\chi(g)$. In particular, $\chi(g) = 0$ if g is self-paired.

3.3 Realisation Part I

In this section we construct the groups $\bar{G}^{(\ell)}$ as subgroups of Sym_n where $n = 24/(\ell-1)$. This construction will be nested in the sense that each group $\bar{G}^{(\ell)}$ will be realised as a subgroup of some $\bar{G}^{(\ell')}$ for $\ell'-1$ a divisor of $\ell-1$.

For $\ell = 2$ let $\Omega^{(2)}$ be a set with 24 elements and let $\mathcal{G} \subset \mathcal{P}(\Omega^{(2)})$ be a copy of the Golay code on $\Omega^{(2)}$. Then the subgroup of the symmetric group $Sym_{\Omega^{(2)}}$ of permutations of the set $\Omega^{(2)}$ that preserves $\mathcal{G}$ is isomorphic to M_{24}. So we may set

$$\bar{G}^{(2)} = \left\{\sigma \in Sym_{\Omega^{(2)}} \mid \sigma(C) \in \mathcal{G}, \forall C \in \mathcal{G}\right\}.$$

For $\ell = 3$ the largest proper divisor of $\ell - 1 = 2$ is $1 = 2-1$. Choose a partition of $\Omega^{(2)}$ into $2/1 = 2$ subsets of $24/2 = 12$ elements such that each 12-element set belongs to $\mathcal{G}$ and denote it $P^{(3)} = \{\Omega^{(3)}, \Omega^{(3)'}\}$. Equivalently, take $\Omega^{(3)}$ to be the symmetric difference of the sets of fixed-points of elements $\sigma, \sigma' \in \bar{G}^{(2)}$ of cycle shape $1^8 2^8$ such that $\sigma\sigma'$ has order 6. (That is, $\Omega^{(3)}$ consists of the points that are fixed by σ or σ' but not both. In turns out that if the product $\sigma\sigma'$ has order 6 then its cycle shape is $1^2 2^2 3^2 6^2$ and so the fixed-point sets of σ and σ' have intersection of size 2 and $\Omega^{(3)}$ and $\Omega^{(3)'} = \Omega^{(2)} \setminus \Omega^{(3)}$ both have 12 elements.) Then the group of permutations of $\Omega^{(3)}$ induced by the subgroup of $\bar{G}^{(2)}$ that fixes this partition is a copy of $\bar{G}^{(3)} \simeq M_{12}$. Explicitly, and to explain what we mean here by fix, if $\bar{G}^{(2)}_{P^{(3)}}$ denotes the subgroup of $\bar{G}^{(2)}$ that fixes the partition $P^{(3)}$ in the sense that

$$\bar{G}^{(2)}_{P^{(3)}} = \left\{ \sigma \in \bar{G}^{(2)} \mid \sigma(C) = C,\, \forall C \in P^{(3)} \right\}, \tag{3.7}$$

and if $\varrho : \bar{G}^{(2)}_{P^{(3)}} \to Sym_{\Omega^{(3)}}$ denotes the natural map obtained by restricting elements of $\bar{G}^{(2)}_{P^{(3)}}$ to $\Omega^{(3)}$,

$$\begin{aligned} \varrho : \bar{G}^{(2)}_{P^{(3)}} &\to Sym_{\Omega^{(3)}} \\ \sigma &\mapsto \sigma|_{\Omega^{(3)}}, \end{aligned} \tag{3.8}$$

then $\bar{G}^{(3)}$ is the image of ϱ. The kernel of this map is the subgroup

$$\bar{G}^{(2)}_{P^{(3)},\Omega^{(3)}} = \left\{ \sigma \in \bar{G}^{(2)}_{P^{(3)}} \mid \sigma(x) = x,\, \forall x \in \Omega^{(3)} \right\} \tag{3.9}$$

comprised of permutations in $\bar{G}^{(2)}_{P^{(3)}}$ that fix every element of $\Omega^{(3)}$ so we have

$$\bar{G}^{(3)} = \varrho\left(\bar{G}^{(2)}_{P^{(3)}}\right) \simeq \bar{G}^{(2)}_{P^{(3)}} / \bar{G}^{(2)}_{P^{(3)},\Omega^{(3)}} \simeq M_{12}.$$

For $\ell = 4$ the largest proper divisor of $\ell - 1 = 3$ is $1 = 2 - 1$. Choose a partition of $\Omega^{(2)}$ into $3/1 = 3$ subsets of $24/3 = 8$ elements that belong to $\mathcal{G}$ and denote it $P^{(4)} = \{\Omega^{(4)}, \Omega^{(4)'}, \Omega^{(4)''}\}$. Equivalently, choose $\Omega^{(4)}$ and $\Omega^{(4)'}$ to be the fixed-point sets of respective elements σ and σ' of order 2 in $\bar{G}^{(2)}$ having cycle shapes $1^8 2^8$ and disjoint fixed-point sets and the property that $\sigma\sigma'$ has order 3. Then the group of permutations of $\Omega^{(4)}$ induced by the subgroup of $\bar{G}^{(2)}$ that fixes this partition is a copy of the group $\bar{G}^{(4)} \simeq AGL_3(2)$ and we have

$$\bar{G}^{(4)} = \varrho\left(\bar{G}^{(2)}_{P^{(4)}}\right) \simeq \bar{G}^{(2)}_{P^{(4)}} / \bar{G}^{(2)}_{P^{(4)},\Omega^{(4)}} \simeq AGL_3(2)$$

where $\bar{G}^{(2)}_{P^{(4)}}$, the map ϱ and the subgroup $\bar{G}^{(2)}_{P^{(4)},\Omega^{(4)}}$ are defined according to the natural analogues of (3.7), (3.8) and (3.9), respectively.

For $\ell = 5$ the largest proper divisor of $\ell - 1 = 4$ is $2 = 3 - 1$. Choose a partition $P^{(5)} = \{\Omega^{(5)}, \Omega^{(5)'}\}$ of $\Omega^{(3)}$ into $4/2 = 2$ subsets of $12/2 = 6$ elements such that neither of the sets in $P^{(5)}$ is contained in a set of size 8 in $\mathcal{G}$. Equivalently, take $\Omega^{(5)}$ to be the symmetric difference of the sets of fixed-points of elements $\sigma, \sigma' \in \bar{G}^{(3)}$ of cycle shape $1^4 2^4$ such that $\sigma\sigma'$ has order 6. (This condition forces the fixed-point sets of σ and σ' to have a single point of intersection so that $\Omega^{(5)}$ and $\Omega^{(5)'} = \Omega^{(3)} \setminus \Omega^{(5)}$ both have 6 elements.) Then the group of permutations of

$\Omega^{(5)}$ induced by the subgroup of $\bar{G}^{(3)}$ that fixes this partition is a copy of $\bar{G}^{(5)} \simeq PGL_2(5)$.

$$\bar{G}^{(5)} = \varrho\left(\bar{G}^{(3)}_{P^{(5)}}\right) \simeq \bar{G}^{(3)}_{P^{(5)}}/\bar{G}^{(3)}_{P^{(5)},\Omega^{(5)}} \simeq PGL_2(5)$$

For $\ell = 7$ the largest proper divisor of $\ell - 1 = 6$ is $3 = 4 - 1$. Choose a partition $P^{(7)} = \{\Omega^{(7)}, \Omega^{(7)'}\}$ of $\Omega^{(4)}$ into $6/3 = 2$ subsets of $8/2 = 4$ elements such that $\Omega^{(7)}$ (and therefore also $\Omega^{(7)'} = \Omega^{(4)} \setminus \Omega^{(7)}$) is the set of fixed points of an element of order 2 in $\bar{G}^{(4)}$ having cycle shape $1^4 2^2$. (There are three conjugacy classes of elements of order 2 in $\bar{G}^{(4)}$ but for just one of these classes do the elements have the cycle shape $1^4 2^2$.) Then the group of even permutations of $\Omega^{(7)}$ induced by the subgroup of $\bar{G}^{(4)}$ that fixes this partition is a copy of the group $\bar{G}^{(7)} \simeq L_2(3)$, which is of course isomorphic to the alternating group on 4 points.

$$\bar{G}^{(7)} = \varrho\left(\bar{G}^{(4)}_{P^{(7)}}\right) \cap Alt_{\Omega^{(7)}} \simeq L_2(3)$$

(The quotient $\bar{G}^{(4)}_{P^{(7)}}/\bar{G}^{(4)}_{P^{(7)},\Omega^{(7)}}$ is isomorphic to the full symmetric group on 4 points so the restriction to even permutations is not redundant.)

The next largest divisor of $7 - 1 = 6$ is $2 = 3 - 1$. An alternative construction of $\bar{G}^{(7)}$ is obtained by choosing a partition $P^{(7)} = \{\Omega^{(7)}, \Omega^{(7)'}, \Omega^{(7)''}\}$ of $\Omega^{(3)}$ into $6/2 = 3$ subsets of $12/3 = 4$ elements such that $\Omega^{(7)}$ and $\Omega^{(7)'}$ are the fixed-point sets of respective elements σ and σ' of order 2 in $\bar{G}^{(3)}$ having cycle shapes $1^4 2^4$ and disjoint fixed-point sets and the property that $\sigma\sigma'$ has order 3. Then the group of permutations of $\Omega^{(7)}$ induced by the subgroup of $\bar{G}^{(3)}$ that fixes this partition is again a copy of the group $\bar{G}^{(7)} \simeq L_2(3)$.

$$\bar{G}^{(7)} = \varrho\left(\bar{G}^{(3)}_{P^{(7)}}\right) \simeq \bar{G}^{(3)}_{P^{(7)}}/\bar{G}^{(3)}_{P^{(7)},\Omega^{(7)}} \simeq L_2(3)$$

For $\ell = 13$ the largest proper divisor of $\ell - 1 = 12$ is $6 = 7 - 1$. Choose a partition $P^{(13)} = \{\Omega^{(13)}, \Omega^{(13)'}\}$ of $\Omega^{(7)}$ into $12/6 = 2$ subsets of $4/2 = 2$ elements such that $\Omega^{(13)}$ is the orbit of an element of order 2 in $\bar{G}^{(7)}$. (That is, choose any partition of $\Omega^{(7)}$ into subsets of size 2. The elements of order 2 in $\bar{G}^{(7)}$ all have cycle shape 2^2.) Then the group of permutations of $\Omega^{(13)}$ induced by the subgroup of $\bar{G}^{(7)}$ that fixes this partition is a copy of the group $\bar{G}^{(13)} \simeq Sym_2$.

$$\bar{G}^{(13)} = \varrho\left(\bar{G}^{(7)}_{P^{(13)}}\right) \simeq \bar{G}^{(7)}_{P^{(13)}}/\bar{G}^{(7)}_{P^{(13)},\Omega^{(13)}} \simeq Sym_2$$

The next largest divisor of $13 - 1 = 12$ is $4 = 5 - 1$. An alternative construction of $\bar{G}^{(13)}$ is obtained by choosing a partition $P^{(13)} = \{\Omega^{(13)}, \Omega^{(13)'}, \Omega^{(13)''}\}$ of $\Omega^{(5)}$ into $12/4 = 3$ subsets of $6/3 = 2$ elements such that $\Omega^{(13)}$ and $\Omega^{(13)'}$ are the fixed-point sets of respective elements σ

and σ' of order 2 in $\bar{G}^{(3)}$ having cycle shapes $1^2 2^2$ and disjoint fixed-point sets and the property that $\sigma\sigma'$ has order 3. Then the group of permutations of $\Omega^{(13)}$ induced by the subgroup of $\bar{G}^{(5)}$ that fixes this partition is a copy of the group $\bar{G}^{(13)} \simeq Sym_2$.

$$\bar{G}^{(13)} = \varrho\left(\bar{G}^{(5)}_{P^{(13)}}\right) \simeq \bar{G}^{(5)}_{P^{(13)}} / \bar{G}^{(5)}_{P^{(13)},\Omega^{(13)}} \simeq Sym_2$$

Remark 3.1. The group $\bar{G}^{(\ell')}_{P^{(\ell)},\Omega^{(\ell)}}$ is trivial in every instance except for when (ℓ',ℓ) is $(2,4)$ or $(4,7)$ so apart from these cases we have $\bar{G}^{(\ell)} \simeq \bar{G}^{(\ell')}_{P^{(\ell)}}$. The group $\bar{G}^{(2)}_{P^{(4)},\Omega^{(4)}}$ has order 8 and $\bar{G}^{(4)}_{P^{(7)},\Omega^{(7)}}$ has order 4.

Remark 3.2. A *Steiner system* with parameters (t,k,n) is the data of an n-element set Ω together with a collection of subsets of Ω of size k, called the *blocks* of the system, with the property that any t-element subset of Ω is contained in a unique block. It is well-known that $M_{24} \simeq \bar{G}^{(2)}$ may be realised as the automorphism group of a Steiner system with parameters $(5,8,24)$ and $M_{12} \simeq \bar{G}^{(3)}$ may be described as the automorphism group of a Steiner system with parameters $(5,6,12)$. Indeed, for $\bar{G}^{(2)} \simeq M_{24}$ we may take the blocks to be the $\binom{24}{5}/\binom{8}{5} = 759$ sets of size 8 in a copy $\mathcal{G}$ of the Golay code on the 24 element set $\Omega = \Omega^{(2)}$ and for $\bar{G}^{(3)} \simeq M_{12}$ regarded as the subgroup of $\bar{G}^{(2)}$ preserving a set $\Omega^{(3)}$ of size 12 in $\mathcal{G}$ we may take the blocks to be the $\binom{12}{5}/\binom{6}{5} = 132$ subsets of size 6 that arise as an intersection $C \cap \Omega^{(3)}$ for $C \in \mathcal{G}$ having size 8. The group $\bar{G}^{(4)} \simeq AGL_3(2)$ also admits such a description: it is the automorphism group of the unique Steiner system with parameters $(3,4,8)$. The blocks of such a system constitute the $\binom{8}{3}/\binom{4}{3} = 14$ codewords of weight 4 in the unique doubly even linear binary code of length 8, the length 8 *Hamming code*.

Remark 3.3. Following [92, Chp. 11] let n be a divisor of 12, take n cards labelled by 0 through $n-1$ and consider the group $M_n < Sym_n$ generated by the *reverse shuffle* $r_n : t \mapsto n-1-t$ and the *Mongean shuffle* $s_n : t \mapsto \min\{2t, 2n-1-2t\}$. Then the notation is consistent with that used above, for the group M_{12} just defined is indeed isomorphic to the Mathieu group of permutations on 12 points. So we have $M_{12} \simeq \bar{G}^{(3)}$ according to Table 1 and it turns out that in fact $M_n \simeq \bar{G}^{(\ell)}$ whenever $n = 24/(\ell-1)$, and so this shuffle construction recovers all of the $\bar{G}^{(\ell)}$ except for $G^{(2)} \simeq M_{24}$ and $G^{(4)} \simeq AGL_3(2)$ (i.e. all $\bar{G}^{(\ell)}$ for ℓ odd).

3.4 Realisation Part II

Having constructed the $\bar{G}^{(\ell)}$ as permutation groups we now describe the umbral groups $G^{(\ell)}$ as groups of signed permutations. We retain the notation of §3.3.

To construct $G^{(3)}$ choose an element τ of order 11 in $\bar{G}^{(3)}$ and (re)label the elements of $\Omega^{(3)}$

so that

$$\Omega^{(3)} = \{\infty, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, X\}$$

and the action of τ on $\Omega^{(3)}$ is given by $x \mapsto x + 1$ modulo 11 (where $X = 10$). Then for any set $C \subset \Omega^{(3)}$ of size 4 there is a unique element $\sigma \in \bar{G}^{(3)}$ with cycle shape $1^4 2^4$ such that the fixed-point set of σ is precisely C. In case $C = \{0, 1, 4, 9\}$ for example $\sigma = (\infty 6)(2X)(35)(78)$ and $\bar{G}^{(3)}$ is generated by σ and τ. Let $\hat{\Omega}^{(3)} = \{e_x \mid x \in \Omega^{(3)}\}$ be a basis for a 12-dimensional vector space over $\mathbb{C}$ say, let $\hat{\tau}$ be the element of the hyperoctahedral group $Oct_{\hat{\Omega}^{(3)}} \simeq Oct_{12}$ given by $\hat{\tau} : e_x \mapsto e_{x+1}$, so that $\hat{\tau} = (0123456789X)$ in (signed) cycle notation (cf. §3.2), and set $\hat{\sigma} = (\infty 6)(\bar{2}\bar{X})(35)(\bar{7}\bar{8})$. Then for $G^{(3)}$ the group generated by $\hat{\sigma}$ and $\hat{\tau}$ we have that $G^{(3)} \simeq 2.M_{12}$ and the image of $G^{(3)}$ under the map $Oct_{\hat{\Omega}^{(3)}} \to S_{\Omega^{(3)}}$ is $\bar{G}^{(3)}$.

To construct $G^{(4)}$ choose an element τ of order 7 in $\bar{G}^{(4)}$ and (re)label the elements of $\Omega^{(4)}$ so that

$$\Omega^{(4)} = \{\infty, 0, 1, 2, 3, 4, 5, 6\}$$

and the action of τ on $\Omega^{(4)}$ is given by $x \mapsto x + 1$ modulo 7. There is a unique conjugacy class of elements of $\bar{G}^{(4)}$ having cycle shape $1^4 2^2$ and the fixed-point sets of these elements are the 14 subsets of $\Omega^{(4)}$ of size 4 comprising the blocks of a $(3, 4, 8)$ Steiner system preserved by $\bar{G}^{(4)}$ (cf. Remark 3.2). The particular block $C = \{0, 1, 3, 6\}$ is the unique one with the property that σ and τ generate $\bar{G}^{(4)}$ in case σ is any (of the 3) involution(s) with C as its fixed-point set. Take $\sigma = (\infty 5)(24)$ and let $\hat{\Omega}^{(4)} = \{e_x \mid x \in \Omega^{(4)}\}$ be a basis for an 8-dimensional vector space over $\mathbb{C}$ say, let $\hat{\tau}$ be the element of the hyperoctahedral group $Oct_{\hat{\Omega}^{(4)}} \simeq Oct_8$ given by $\hat{\tau} : e_x \mapsto e_{x+1}$, so that $\hat{\tau} = (0123456)$, and set $\hat{\sigma} = (\infty 5)(\bar{0})(\bar{1})(24)$. Then for $G^{(4)}$ the group generated by $\hat{\sigma}$ and $\hat{\tau}$ we have that $G^{(4)} \simeq 2.AGL_3(2)$ and the image of $G^{(4)}$ under the map $Oct_{\hat{\Omega}^{(4)}} \to S_{\Omega^{(4)}}$ is $\bar{G}^{(4)}$.

For the remaining $\ell \geq 5$ we can construct $G^{(\ell)}$ from $G^{(3)}$ by proceeding in analogy with the constructions of $\bar{G}^{(\ell)}$ from $\bar{G}^{(3)}$ (and its subgroups) given earlier in this section. Concretely, for $\ell = 5$ we may set

$$G^{(5)} = \hat{\varrho}\left(G^{(3)}_{\hat{P}^{(5)}}\right) \simeq G^{(3)}_{\hat{P}^{(5)}}/G^{(3)}_{\hat{P}^{(5)},\hat{\Omega}^{(5)}} \simeq GL_2(5)/2$$

where $\hat{\Omega}^{(5)} = \{e_x \mid x \in \Omega^{(5)}\}$ is a basis for $24/(5-1) = 6$-dimensional vector space over $\mathbb{C}$, we define $\hat{C} = \{e_x \mid x \in C\}$ for $C \subset \Omega^{(5)}$ and set $\hat{P}^{(5)} = \{\hat{\Omega}^{(5)}, \hat{\Omega}^{(5)\prime}\}$ for $P^{(5)} = \{\Omega^{(5)}, \Omega^{(5)\prime}\}$ as in the construction of $\bar{G}^{(5)}$ given above, the groups $G^{(3)}_{\hat{P}^{(5)}}$ and $G^{(3)}_{\hat{P}^{(5)},\hat{\Omega}^{(5)}}$ are defined by

$$G^{(3)}_{\hat{P}^{(5)}} = \left\{\sigma \in G^{(3)} \mid \sigma(\operatorname{Span}\hat{C}) \subset \operatorname{Span}\hat{C},\, \forall \hat{C} \in \hat{P}^{(5)}\right\}, \tag{3.10}$$

$$G^{(3)}_{\hat{P}^{(5)},\hat{\Omega}^{(5)}} = \left\{\sigma \in G^{(3)}_{\hat{P}^{(5)}} \mid \sigma(e_x) = e_x,\, \forall x \in \Omega^{(5)}\right\}, \tag{3.11}$$

and $\hat{\varrho}$ is the map $G^{(3)}_{\hat{P}^{(5)}} \to Oct_{\hat{\Omega}^{(5)}}$ obtained by restriction.

$$\begin{aligned} \hat{\varrho} : G^{(3)}_{\hat{P}^{(5)}} &\to Oct_{\hat{\Omega}^{(5)}} \\ \sigma &\mapsto \sigma|_{\operatorname{Span}\hat{\Omega}^{(5)}} \end{aligned} \tag{3.12}$$

In direct analogy with this we have that

$$\begin{aligned} G^{(7)} &= \hat{\varrho}\left(G^{(3)}_{\hat{P}^{(7)}}\right) \simeq G^{(3)}_{\hat{P}^{(7)}}/G^{(3)}_{\hat{P}^{(7)},\hat{\Omega}^{(7)}} \simeq SL_2(3), \\ G^{(13)} &= \hat{\varrho}\left(G^{(5)}_{\hat{P}^{(13)}}\right) \simeq G^{(5)}_{\hat{P}^{(13)}}/G^{(5)}_{\hat{P}^{(13)},\hat{\Omega}^{(13)}} \simeq 4, \end{aligned}$$

when $G^{(\ell')}_{\hat{P}^{(\ell)}}$ and $G^{(\ell')}_{\hat{P}^{(\ell)},\hat{\Omega}^{(\ell)}}$ are defined as in (3.10) and (3.11), respectively, for $P^{(\ell)} = \{\Omega^{(\ell)}, \cdots\}$ as specified in the construction of $\bar{G}^{(\ell')}$ given above, and $\hat{\varrho}$ as in (3.12).

We conclude this section with explicit signed permutation presentations of the $G^{(\ell)}$ as subgroups of the $Oct_{\hat{\Omega}^{(\ell)}}$ for $\ell \geq 3$. The presentations for $\ell = 3$ and $\ell = 4$ were obtained above, and the remaining ones can be found in a similar manner. In each case we label the index set $\Omega^{(\ell)}$ so that $\Omega^{(\ell)} = \{\infty, 0, \cdots, n-1\}$ where $n = 24/(\ell-1) - 1 = (25-\ell)/(\ell-1)$ and seek a presentation for which the coordinate permutation $e_x \mapsto e_{x+1}$ (with indices read modulo n) is an element of order n in $G^{(\ell)}$ (although we must take this element to be trivial in case $\ell = 13$).

$$G^{(3)} = \left\langle(\infty 6)(\bar{2}\bar{X})(35)(\bar{7}\bar{8}), (0123456789X)\right\rangle \tag{3.13}$$

$$G^{(4)} = \langle(\infty 5)(\bar{0})(\bar{1})(24), (0123456)\rangle \tag{3.14}$$

$$G^{(5)} = \langle(\infty\bar{0})(3\bar{1})(2\bar{4}), (01234)\rangle \tag{3.15}$$

$$G^{(7)} = \langle(\infty\bar{0})(\bar{1}2), (012)\rangle \tag{3.16}$$

$$G^{(13)} = \langle(\infty\bar{0})\rangle \tag{3.17}$$

Equipped now with explicit realisations $G^{(\ell)} < Oct_n$ of the umbral groups $G^{(\ell)}$ as signed permutation groups we define symbols $\Pi^{(\ell)}_g$, $\chi^{(\ell)}_g$, $\bar{\Pi}^{(\ell)}_g$ and $\bar{\chi}^{(\ell)}_g$ as follows for $\ell \in \Lambda$ and $g \in G^{(\ell)}$. We write $\Pi^{(\ell)}_g$ for the signed permutation Frame shape attached to $g \in G^{(\ell)}$ (when regarded as an element of Oct_n) as defined in §3.2 and we write $\bar{\Pi}^{(\ell)}_g$ for the cycle shape attached to the image of $g \in G^{(\ell)}$ under the composition $G^{(\ell)} \to Oct_n \to Sym_n$. We define $g \mapsto \chi_g^{(\ell)}$ by restricting the signed permutation character (cf. (3.3)) to $G^{(\ell)}$ and we define $g \mapsto \bar{\chi}_g^{(\ell)}$ by restricting the unsigned permutation character (cf. (3.4)) to $G^{(\ell)}$. We call $\chi_g^{(\ell)}$ the *signed twisted Euler character* attached to $g \in G^{(\ell)}$ and we call $\bar{\chi}_g^{(\ell)}$ the *unsigned twisted Euler character* attached to $g \in G^{(\ell)}$.

As is explained in §3.2 the data $g \mapsto \Pi_g^{(\ell)}$ is sufficient to determine the cycle shapes $\bar{\Pi}_g^{(\ell)}$ and the twisted Euler characters $\chi_g^{(\ell)}$ and $\bar{\chi}_g^{(\ell)}$. We will attach vector-valued mock modular forms $H_g^{(\ell)}$ to each $g \in G^{(\ell)}$ in §4 and it will develop that the shadows and multiplier systems of these functions are encoded in the Frame shapes $\Pi_g^{(\ell)}$. We list the Frame shapes $\Pi_g^{(\ell)}$ and $\bar{\Pi}_g^{(\ell)}$ and the twisted Euler characters $\chi_g^{(\ell)}$ and $\bar{\chi}_g^{(\ell)}$ for all $g \in G^{(\ell)}$ and $\ell \in \Lambda$ in the tables of §B.2.

We conclude this section with an extraordinary property relating the Frame shapes $\Pi_g^{(\ell)}$ and $\bar{\Pi}_g^{(\ell)}$ attached to the umbral groups $G^{(\ell)}$ at $\ell = 2$ and $\ell = 4$ which can be verified explicitly using the tables of §B.2.

Proposition 3.4. *Let $g \in G^{(4)}$ and suppose that the Frame shape $\Pi_g^{(4)} = \prod_k k^{m_g(k)}$ is a cycle shape, so that $m_g(k) \geq 0$ for all k. Then there exists $g' \in G^{(2)}$ with $o(g') = 2o(g)$ such that*

$$\bar{\Pi}_{g'}^{(2)} = \prod_k k^{\bar{m}_g(k)} (2k)^{m_g(k)} \tag{3.18}$$

when $\bar{\Pi}_g^{(4)} = \prod_k k^{\bar{m}_g(k)}$ except in case g is of class 4B.

As will be explained in §4 this result implies direct relationships between the functions $H_{g'}^{(2)}$ and $H_g^{(4)}$ when g and g' are as in the statement of the proposition.

3.5 Dynkin Diagrams Part I

The *McKay correspondence* [93] relates finite subgroups of $SU(2)$ to the affine Dynkin diagrams of ADE type by associating irreducible representations of the finite groups to nodes of the corresponding diagrams and by now is well understood in terms of resolutions of *simple singularities* $\mathbb{C}^2/G$ for $G < SU(2)$ [94, 95]. A more mysterious Dynkin diagram correspondence also due to McKay is his *monstrous E_8 observation* [93] (see also [96, §14]) which associates nine of the conjugacy classes of the monster group to the nodes of the affine E_8 Dynkin diagram and extends to similar correspondences relating certain subgroups of the Monster to other affine Dynkin diagrams [97]. We find analogues of both McKay's Dynkin diagram observations manifesting in the groups $G^{(\ell)}$, as we shall now explain.

For $\ell \in \Lambda = \{2, 3, 4, 5, 7, 13\}$ the number $(25 - \ell)/(\ell - 1)$ is an odd integer p such that $p + 1$ divides 24 and is a prime in case ℓ is not 13. Recall the construction of $\bar{G}^{(\ell)}$ as permutations of

$\Omega^{(\ell)}$ from §3.3. By inspection, with the assistance of [98], we obtain the following lemma.

Lemma 3.5. *Let $\ell \in \{2,3,4,5,7\}$ and set $p = (25-\ell)/(\ell-1)$. Then the group $\bar{G}^{(\ell)}$ has a unique conjugacy class of subgroups isomorphic to $L_2(p)$ that act transitively on the degree $p+1 = 24/(\ell-1)$ permutation representation of $\bar{G}^{(\ell)}$ on $\Omega^{(\ell)}$.*

Armed with Lemma 3.5 we pick a transitive subgroup of $\bar{G}^{(\ell)}$ isomorphic to $L_2(p)$ for each ℓ in $\{2,3,4,5,7\}$—these being the cases that $(25-\ell)/(\ell-1)$ is prime—and denote it $\bar{L}^{(\ell)}$. For future reference we note that there are two conjugacy classes of subgroups of $\bar{G}^{(\ell)}$ isomorphic to $L_2(p)$ in case $\ell \in \{3,4\}$ but in each case only one of these classes contains subgroups acting transitively.

Lemma 3.6. *For $\ell \in \{3,4\}$ and $p = (25-\ell)/(\ell-1)$ there is a unique conjugacy class of subgroups of $\bar{G}^{(\ell)}$ isomorphic to $L_2(p)$ that does not act transitively in the degree $p+1 = 24/(\ell-1)$ permutation representation of $\bar{G}^{(\ell)}$ on the set $\Omega^{(\ell)}$.*

It is a result due to Galois (proven in a letter to Chevalier written on the eve of his deadly duel [92, Chp. 10]) that the group $L_2(p)$ has no transitive permutation representation of degree less than $p+1$ if $p > 11$. However for the remaining primes p not exceeding 11 there are transitive permutation representations on exactly p points, and in fact we have the following statement.

Lemma 3.7. *Let $p \in \{2,3,5,7,11\}$ and set $\bar{L} = L_2(p)$. Then there is a subgroup $\bar{D} < \bar{L}$ of index p with the property that $\bar{L} = \langle\sigma\rangle\bar{D}$ for σ an element of order p in $\bar{L}$, so that every element $g \in \bar{L}$ admits a unique expression $g = sd$ where $s \in \langle\sigma\rangle$ and $d \in \bar{D}$.*

Remark 3.8. See [92, Chp. 10] for applications of the result of Lemma 3.7 to exceptional isomorphisms among finite simple groups of small order, and see [99, 100] for an application of the case that $p = 11$ to the truncated icosahedron (which describes the structure of buckminsterfullerenes, a.k.a. buckyballs).

According to Lemma 3.7 we obtain a permutation representation of degree p for $\bar{L}$ by taking the natural action of $\bar{L}$ on cosets of $\bar{D}$ and this action is transitive since σ induces a p-cycle. Using this result we choose a copy of $\bar{D}$ in $\bar{L} = \bar{L}^{(\ell)}$ for each $\ell \in \{3,4,5,7\}$ and denote it $\bar{D}^{(\ell)}$. (Such subgroups are uniquely determined up to isomorphism, but there are as many conjugacy classes of subgroups of $\bar{L}$ isomorphic to $\bar{D}$ as there are conjugacy classes of elements of order p in $\bar{L}$, which is to say there are 2 classes in case $p \in \{11,7\}$ and a unique class in case $p \in \{5,3\}$.) We describe the groups $\bar{D}^{(\ell)}$ in Table 2 and observe that each one is isomorphic to a finite group $\bar{D}_0 < SO_3(\mathbb{R})$. As such there is a corresponding finite group $D_0 < SU(2)$ that maps onto $\bar{D}_0$

Table 2: The Umbral Groups and Dynkin Data

ℓ	2	3	4	5	7	13
p	23	11	7	5	3	1
$G^{(\ell)}$	M_{24}	$2.M_{12}$	$2.AGL_3(2)$	$GL_2(5)/2$	$SL_2(3)$	4
$L^{(\ell)}$		$SL_2(11)$	$SL_2(7)$		$SL_2(3)$	
$D^{(\ell)}$		$2.Alt_5$	$2.Sym_4$		Q_8	
$\bar{G}^{(\ell)}$	M_{24}	M_{12}	$AGL_3(2)$	$PGL_2(5)$	$L_2(3)$	2
$\bar{L}^{(\ell)}$	$L_2(23)$	$L_2(11)$	$L_2(7)$	$L_2(5)$	$L_2(3)$	
$\bar{D}^{(\ell)}$		Alt_5	Sym_4	Alt_4	2^2	
$\Delta^{(\ell)}$		$\hat{E}_8$	$\hat{E}_7$	$\hat{E}_6$	$\hat{D}_4$	

via the 2-fold covering $SU(2) \to SO_3(\mathbb{R})$, and a corresponding Dynkin diagram $\Delta^{(\ell)}$ according to the McKay correspondence. We list the Dynkin diagrams $\Delta^{(\ell)}$ also in Table 2.

Traditionally finite subgroups of $SO_3(\mathbb{R})$ are called *ternary*, owing to the fact that their elements are described using 3×3 matrices, and finite subgroups of $SU(2)$ are called *binary*. The map $SU(2) \to SO_3(\mathbb{R})$ determines a correspondence between binary and ternary groups whereby the ternary polyhedral groups (of orientation preserving symmetries of regular polyhedra) correspond to binary groups of the form $2.Alt_5$, $2.Sym_4$ and $2.Alt_4$ (depending upon the polyhedron), a dihedral subgroup $Dih_n < SO_3(\mathbb{R})$ corresponds to a generalised quaternion group of order $4n$ in $SU(2)$, and a binary cyclic group corresponds to a ternary cyclic group of the same order. Thus, given a finite subgroup $\bar{D}_0 < SO_3(\mathbb{R})$, we may speak of the *associated binary group* $D_0 < SU(2)$.

Next we observe that each group $G^{(\ell)}$ contains a subgroup $D^{(\ell)}$ isomorphic to the binary group associated to $\bar{D}^{(\ell)}$ (when $\bar{D}^{(\ell)}$ is regarded as a subgroup of $SO_3(\mathbb{R})$) for all of the above cases except when $\ell = p = 5$. More particularly, for $\ell \in \{3, 7\}$ there is a unique conjugacy class of subgroups isomorphic to the binary group attached to $\bar{D}^{(\ell)}$ while for $\ell = 4$ there are two such conjugacy classes and for $\ell = 5$ there are none. For each $\ell \in \{3, 4, 7\}$ we pick a subgroup isomorphic to the binary group attached to $\bar{D}^{(\ell)}$ and denote it $D^{(\ell)}$ and we display the (isomorphism types of the) groups $D^{(\ell)}$ in Table 2. We write Q_8 there for the quaternion group of order 8.

To see how the $D^{(\ell)}$ arise in $G^{(\ell)}$ (and fail to do so in the case that $\ell = 5$) recall from Lemma 3.5 that $\bar{G}^{(\ell)}$ contains a unique transitive subgroup isomorphic to $L_2(p)$ up to conjugacy for $\ell \in \{3, 4, 5, 7\}$. The preimage of such a subgroup under the natural map $G^{(\ell)} \to \bar{G}^{(\ell)}$ (cf. §3.4) is a double cover of $L_2(p)$ that is in fact isomorphic to $SL_2(p)$ (cf. §3.1) except when $\ell = 5$. In case $\ell = 5$ we have $G^{(5)} \simeq GL_2(5)/2$ which has the same order as $SL_2(5)$ but is not isomorphic to it. We see then that for $\ell \in \{3,4,7\}$ we may find a copy of $SL_2(p)$ in the preimage of $\bar{L}^{(\ell)}$ under the map $G^{(\ell)} \to \bar{G}^{(\ell)}$; we do so and denote it $L^{(\ell)}$. Then we may take $D^{(\ell)}$ to be the preimage of $\bar{D}^{(\ell)}$ under the map $L^{(\ell)} \to \bar{L}^{(\ell)}$. The fact that there is no $SL_2(5)$ in $G^{(5)}$ explains why there is no group $D^{(5)}$.

Observe that the rank of $\Delta^{(\ell)}$ is given by $11-\ell$ for each $\ell \in \{3,4,5,7\}$. This may be taken as a hint to the following uniform construction of the $\Delta^{(\ell)}$. Starting with the (finite type) E_8 Dynkin diagram, being star shaped with three *branches*, construct a sequence of diagrams iteratively by removing the end node from a branch of maximal length at each iteration. In this way we obtain the sequence E_8, E_7, E_6, D_5, D_4, A_3, A_2, A_1, and it is striking to observe that our list $\Delta^{(\ell)}$, obtained by applying the McKay correspondence to distinguished subgroups of the $G^{(\ell)}$, is a subsequence of the one obtained from this by affinisation.

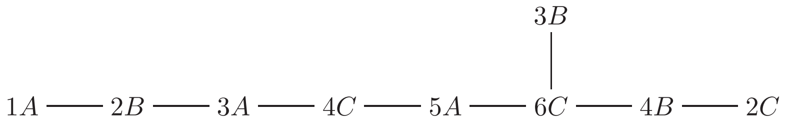
We summarise the main observations of this section as follows.

> *For $\ell \in \{3,4,5,7\}$ the group $\bar{G}^{(\ell)}$ admits a distinguished isomorphism class of subgroups $\bar{D}^{(\ell)}$. This connects $\bar{G}^{(\ell)}$ to a Dynkin diagram of rank $11-\ell$, for the group $\bar{D}^{(\ell)}$ is the image in $SO_3(\mathbb{R})$ of a finite subgroup $D^{(\ell)}$ in $SU(2)$ which corresponds to the affinisation $\Delta^{(\ell)}$ of a Dynkin diagram of rank $11-\ell$ according to McKay's correspondence. The Dynkin diagrams arising belong to a naturally defined sequence.*

3.6 Dynkin Diagrams Part II

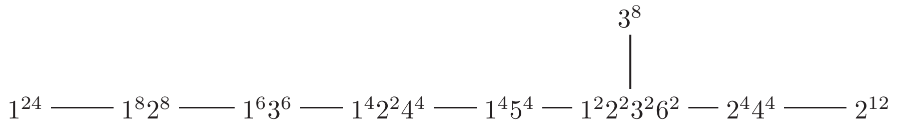
Recall from Lemma 3.6 that the cases $\ell \in \{3,4\}$ are distinguished in that for such ℓ the group $\bar{G}^{(\ell)}$ has a unique conjugacy class of subgroups isomorphic to $L_2(p)$ and not acting transitively in the degree $p+1$ permutation representation (cf. §3.3). Since such an $L_2(p)$ subgroup acts non-trivially on $\Omega^{(\ell)}$ it follows from Lemma 3.7 that it must have one fixed point and act transitively on the remaining p points of $\Omega^{(\ell)}$, and thus we have witnesses within $G^{(\ell)}$ to the exceptional degree p permutation representations of $L_2(p)$ in case $\ell \in \{3,4\}$ and $p \in \{11,7\}$. For these two special cases of lambency 3 and 4 we find direct analogues of McKay's monstrous E_8 observation [93] attaching certain conjugacy classes of $G^{(\ell)}$ to the nodes of $\Delta^{(\ell)}$. At lambency 5 and 7 (where p is 5 and 3, respectively) we find similar analogues where the diagram $\Delta^{(\ell)}$ is replaced by one obtained via folding with respect to a diagram automorphism of order 3.

In case $\ell = 3$ let T denote the conjugacy class of elements g of order 2 in $G^{(3)} \simeq 2.M_{12}$ such that g has 4 fixed points in both the signed and unsigned permutation representations of $G^{(3)}$. This is the conjugacy class labelled $2B$ in Table 9 and we have $\chi_g^{(3)} = \bar{\chi}_g^{(3)} = 4$ in the notation of Table 15. Then $T^2 = \{gh \mid g,h \in T\}$ is a union of conjugacy classes of $G^{(3)}$ and in fact there are exactly nine of the twenty-six conjugacy classes of $G^{(3)}$ that appear. In (3.19) we use these classes (and the notation of Table 9) to label the affine E_8 Dynkin diagram.



(3.19)

Observe that the labelling of (3.19) recovers the highest root labelling when the classes are replaced with their orders. In (3.20) we replace the labels of (3.19) with the cycle shapes attached to these classes via the total permutation action (cf. §3.2) of $G^{(3)}$ on the $24 = 2.24/(\ell - 1)$ elements $\{\pm e_i \mid i \in \Omega^{(3)}\}$ appearing in the signed permutation representation of $G^{(3)}$ (cf. §3.4).



(3.20)

Explicitly, the cycle shape attached to $g \in G^{(3)}$ is the total permutation Frame shape $\tilde{\Pi}_g^{(3)} = \Pi_g^{(3)} \bar{\Pi}_g^{(3)}$ realised as the product (cf. §3.2) of the signed and unsigned permutation Frame shapes attached to $G^{(3)}$ (cf. Table 15).

Remark 3.9. The conjugacy class labelled $2C$ in Table 9 consists of elements of the form gz where g belongs to the class $2B$ (denoted T above) and z is the central involution of $G^{(3)}$, so we obtain exactly the nine conjugacy classes of (3.19) by considering products gh where g and h are involutions in the class $2C$.

Remark 3.10. In [101] an analogue of McKay's monstrous E_8 observation is found for M_{24}. Namely, it is observed that there are exactly nine conjugacy classes of elements of M_{24} having representatives of the form gh where g and h belong to the class labelled 2A in Table 8—this is the class with cycle shape $1^8 2^8$ in the defining degree 24 permutation representation—and the nodes of the affine E_8 Dynkin diagram can be labelled by these nine classes in such a way that their orders recover the highest root labelling. This condition leaves some ambiguity as to the correct placement of the two classes of order 2, and similarly for the pairs of orders 3 and 4, but an analogue of the procedure described in [102] is shown to determine a particular choice that recovers the original correspondence for the monster group via the modular functions of monstrous moonshine. Observe that if we express the $G^{(2)} \simeq M_{24}$ labelling of the affine E_8 diagram given in [101] using cycle shapes to name the conjugacy classes then we recover exactly the labelling (3.20) obtained here from $G^{(3)} \simeq 2.M_{12}$.

In case $\ell = 4$ let T denote the conjugacy class of elements g of order 2 in $G^{(4)}$ such that

g has 4 fixed points in the unsigned permutation representation of $G^{(4)}$ but has 2 fixed points and 2 anti-fixed points in the signed permutation representation. This is the conjugacy class labelled $2C$ in Table 10 and we have $\chi_g^{(4)}=0$ and $\bar{\chi}_g^{(4)}=4$ in the notation of Table 16. Then T^2 is the union of eight of the sixteen conjugacy classes of $G^{(4)}$ and we use these classes (and the notation of Table 10) to label the nodes of the affine E_7 Dynkin diagram in (3.21).

$$\begin{tikzcd}
 & & & 2B & & & \\
1A \arrow[r, dash] & 2C \arrow[r, dash] & 3A \arrow[r, dash] & 4C \arrow[u, dash] \arrow[r, dash] & 6A \arrow[r, dash] & 4A \arrow[r, dash] & 2A
\end{tikzcd} \tag{3.21}$$

In (3.22) we replace the labels of (3.21) with the cycle shapes $\bar{\Pi}_g^{(4)}$ (cf. Table 16) attached to these classes via the degree 8 permutation action of $G^{(4)}$ on $\Omega^{(4)}$. The orders of these permutations are the orders of the images of the corresponding elements of $G^{(4)}$ under the map $G^{(4)} \to \bar{G}^{(4)}$ so the labelling (3.22) demonstrates that we recover the highest root labelling of the affine E_7 Dynkin diagram when we replace the conjugacy classes of (3.21) with the orders of their images in $\bar{G}^{(4)}$.

$$\begin{tikzcd}
 & & & 2^4 & & & \\
1^8 \arrow[r, dash] & 1^4 2^2 \arrow[r, dash] & 1^2 3^2 \arrow[r, dash] & 1^2 2^1 4^1 \arrow[u, dash] \arrow[r, dash] & 1^2 3^2 \arrow[r, dash] & 2^4 \arrow[r, dash] & 1^8
\end{tikzcd} \tag{3.22}$$

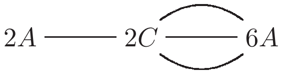
In (3.23) we replace the labels of (3.21) with the cycle shapes of degree 24 given by the products $\Pi_g^{(4)}\bar{\Pi}_g^{(4)}\bar{\Pi}_g^{(4)}$ (cf. Table 16) for g an element of $G^{(4)}$ arising in (3.21). Observe that all the cycle shapes in (3.23) are *balanced*, meaning that they are invariant (and well-defined) under the operation $\prod_{k\geq 1} k^{m(k)} \mapsto \prod_{k\geq 1}(N/k)^{m(k)}$ for some integer $N>1$, and constitute a subset of the cycle shapes attached to $G^{(3)}$ in (3.20). Observe also that the order two symmetry of the affine E_7 diagram that identifies the two long branches is realised by squaring: the cycle shapes obtained by squaring permutations represented by the cycle shapes on the right-hand branch of (3.23) are just those that appear on the left-hand branch.

$$\begin{tikzcd}
 & & & 2^{12} & & & \\
1^{24} \arrow[r, dash] & 1^8 2^8 \arrow[r, dash] & 1^6 3^6 \arrow[r, dash] & 1^4 2^2 4^4 \arrow[u, dash] \arrow[r, dash] & 1^2 2^2 3^2 6^2 \arrow[r, dash] & 2^4 4^4 \arrow[r, dash] & 1^8 2^8
\end{tikzcd} \tag{3.23}$$

Remark 3.11. It is interesting to note that while the conjugacy class $2C$ is stable under multiplication by the central involution there are just eight conjugacy classes of $G^{(4)}$ that are contained in T^2 when T is the class labelled $4A$ in Table 10. In fact the eight conjugacy classes appearing are $8A$ together with all those of (3.21) except for $4C$. Thus we obtain analogues of the labelings (3.21), (3.22) and (3.23) where $4C$, $1^2 2^1 4^1$ and $1^4 2^2 4^4$, respectively, are replaced by $8A$, 4^2 and $4^2 8^2$, respectively. Under this correspondence the highest root labelling is again obtained by taking orders in $\bar{G}^{(4)}$, but while the eta product attached to $4^2 8^2$ is multiplicative it only has weight 2 (i.e. less than 4) and so does not appear in [101].

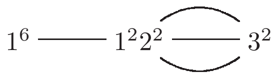
Remark 3.12. The correspondence of [101] may be viewed as attaching the nine multiplicative eta products of weight at least 4 to the nodes of the affine E_8 Dynkin diagram when we regard a cycle shape $\prod_{k\geq 1} k^{m(k)}$ as a shorthand for the *eta product* function $\prod_{k\geq 1} \eta(k\tau)^{m(k)}$ (cf. §A.1). Observe that the cycle shapes appearing in (3.23) are just those whose corresponding eta products are multiplicative, have weight at least 4, and have level dividing 24. (The eta products defined by $1^4 5^4$ and 3^8 have level 5 and 9, respectively.)

For $\ell = 5$ the Dynkin diagram $\Delta^{(5)} = \hat{E}_6$ admits a *Sym*$_3$ group of automorphisms. Folding by either of the three-fold symmetries we obtain the affine Dynkin diagram of type G_2. Let T be either of the two conjugacy class of elements g of order 4 in $G^{(5)}$ such that g^2 is central. Then we have $\chi_g^{(5)} = \bar{\chi}_g^{(5)} = 0$ in the notation of Table 17 and T is either the class labelled $4A$ in Table 11 or the class labelled $4B$, and there are exactly three of the fourteen conjugacy classes of $G^{(5)}$ that are subsets of T^2. In (3.24) we use these classes (and the notation of Table 11) to label the affine G_2 Dynkin diagram.



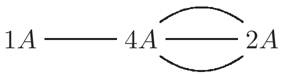
(3.24)

In (3.25) we replace the labels of (3.24) with the cycle shapes $\bar{\Pi}_g^{(5)}$ (cf. Table 17) attached to these classes via the degree 6 permutation action of $G^{(5)}$ on $\Omega^{(5)}$. The orders of these permutations are the orders of the images of the corresponding elements of $G^{(5)}$ under the map $G^{(5)} \to \bar{G}^{(5)}$ so the labelling (3.25) demonstrates that we recover the highest root labelling of the affine G_2 Dynkin diagram—which is the labelling induced form the highest root labelling of $\hat{E}_6$—when we replace the conjugacy classes of (3.24) with the orders of their images in $\bar{G}^{(5)}$.



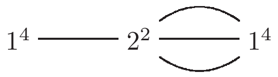
(3.25)

For $\ell = 7$ the Dynkin diagram $\Delta^{(7)} = \hat{D}_4$ admits a *Sym*$_4$ group of automorphisms. Folding by any three-fold symmetry we again obtain the affine Dynkin diagram of type G_2. Let T be the unique conjugacy class of elements of order 4 in $G^{(7)}$. Then we have $\chi_g^{(7)} = \bar{\chi}_g^{(7)} = 0$ in the notation of Table 18 and T is the class labelled $4A$ in Table 12, and there are exactly three of the seven conjugacy classes of $G^{(7)}$ that have a representative of the form gh for some $g, h \in T$ (and T^2 is the union of these three conjugacy classes). In (3.26) we use these classes (and the notation of Table 12) to label the affine G_2 Dynkin diagram.



(3.26)

In (3.27) we replace the labels of (3.26) with the cycle shapes $\bar{\Pi}_g^{(7)}$ (cf. Table 18) attached to these classes via the degree 4 permutation action of $G^{(7)}$ on $\Omega^{(7)}$. The orders of these permutations are the orders of the images of the corresponding elements of $G^{(7)}$ under the map $G^{(7)} \to \bar{G}^{(7)}$ so the labelling (3.27) demonstrates that we recover the labelling of the affine G_2 Dynkin diagram that is induced by the highest root labelling of $\hat{D}_4$ when we replace the conjugacy classes of (3.26) with the orders of their images in $\bar{G}^{(7)}$.



(3.27)

We summarise the results of this section with the following statement.

> *For $\ell \in \{3, 4, 5, 7\}$ there is an analogue of McKay's monstrous E_8 observation that relates $G^{(\ell)}$ to (a folding of) the affinisation of a Dynkin diagram of rank $11 - \ell$, and the diagrams arising are precisely those that appear in §3.5.*

We conclude by mentioning that McKay's monstrous observation has partially been explained, using vertex operator algebra theory, by the work of Sakuma [103] and Lam–Yamada–Yamauchi [104, 105]. Important related work appears in [106, 107].

4 McKay–Thompson Series

In §2.5 we made the observation that the first few positive degree Fourier coefficients of the vector-valued mock modular forms $H^{(\ell)} = \left(H_r^{(\ell)}\right)$ coincide (up to a factor of 2) with dimensions of irreducible representations of the group $G^{(\ell)}$ described in the previous section. This observation suggests the possibility that

$$H_r^{(\ell)}(\tau) = \sum_{k \in \mathbb{Z}} c_r^{(\ell)}(k - r^2/4\ell) q^{k - r^2/4\ell} = -2\delta_{r,1} q^{-1/4\ell} + \sum_{\substack{k \in \mathbb{Z} \\ r^2 - 4k\ell < 0}} \dim K^{(\ell)}_{r, k - r^2/4\ell} q^{k - r^2/4\ell} \tag{4.1}$$

for some $\mathbb{Z} \times \mathbb{Q}$-graded infinite-dimensional $G^{(\ell)}$-module $K^{(\ell)} = \bigoplus_{r,d} K^{(\ell)}_{r,d}$. To further test this possibility we would like to see if there are similar vector-valued mock modular forms

$H_g^{(\ell)} = \big(H_{g,r}^{(\ell)}\big)$ whose positive degree Fourier coefficients recover characters of representations of $G^{(\ell)}$. In other words, for an element g of the group $G^{(\ell)}$ we would like to see if we can find a mock modular form $H_g^{(\ell)}$ compatible with the hypothesis that

$$H_{g,r}^{(\ell)}(\tau) = \sum_k c_{g,r}^{(\ell)}(k - r^2/4\ell)q^{k-r^2/4\ell} = -2\delta_{r,1}q^{-1/4\ell} + \sum_{\substack{k\in\mathbb{Z}\\ r^2-4k\ell<0}} \mathrm{tr}_{K^{(\ell)}_{r,k-r^2/4\ell}}(g)q^{k-r^2/4\ell} \tag{4.2}$$

for a hypothetical bi-graded $G^{(\ell)}$-module $K^{(\ell)}$. We refer to such a generating function as a *McKay–Thompson series*. Notice that we recover the generating functions of the $\dim K_{r,d}^{(\ell)}$ when g is the identity element. Moreover, the McKay–Thompson series attached to $g \in G^{(\ell)}$ is invariant under conjugation, since the trace is such, so $H_g^{(\ell)}$ depends only on the conjugacy class $[g]$ of g.

Having such McKay–Thompson series for each conjugacy class of $G^{(\ell)}$ not only provides strong evidence for the existence of the $G^{(\ell)}$-module $K^{(\ell)}$ but in fact it uniquely specifies it up to $G^{(\ell)}$-module isomorphism. This is because, since there are as many irreducible representations as conjugacy classes of a finite group, given the characters $\mathrm{tr}_{K_{r,d}^{(\ell)}}(g)$ for all $[g] \subset G^{(\ell)}$ we can simply invert the character table to have a unique decomposition of the conjectural module $K_{r,d}^{(\ell)}$ into irreducible representations. What is not clear, a priori, is that we will end up with a decomposition into non-negative integer multiples of irreducible representations of $G^{(\ell)}$, but remarkably it appears that this property holds for all $\ell \in \Lambda$ and all bi-degrees (r, d). For evidence in support of this see the explicit decompositions for small degrees tabulated in §D.

In this section we construct a set of vector-valued mock modular forms $H_g^{(\ell)} = \big(H_{g,r}^{(\ell)}\big)$ for each lambency ℓ and we formulate a precise conjecture implying (4.2) in §5.1.

4.1 Forms of Higher Level

In §2 we have discussed the relation between certain vector-valued mock modular forms, meromorphic Jacobi forms of weight 1 and Jacobi forms of weight 0 under the group $SL_2(\mathbb{Z})$. In order to investigate the McKay–Thompson series of the groups $G^{(\ell)}$ we need to generalise the discussion in §2.5 to modular forms of higher level and consider forms transforming under $\Gamma_0(N)$ (cf. (A.2)) with $N > 1$.

As in §2 we would like to consider $(\ell - 1)$-vector-valued mock modular forms $\big(H_{g,r}^{(\ell)}\big)$ for a group $\Gamma_0(N_g) < SL_2(\mathbb{Z})$ with shadow given by the unary theta series $S^{(\ell)}$ (cf. (2.26)). The levels N_g will be specified explicitly for all g in §4.8. A first difference between the case of $N_g = 1$ and $N_g > 1$ is the following. Since the components of $S^{(\ell)} = \big(S_r^{(\ell)}\big)$ have more than one orbit under $\Gamma_0(N_g)$ when N_g is even it is natural to consider mock modular forms with shadows given by

$\big(\chi_{g,r}^{(\ell)} S_r^{(\ell)}\big)$ where the $\chi_{g,r}^{(\ell)}$ are not necessarily equal for different values of r. It will develop that in the cases of interest to us $\chi_{g,r}^{(\ell)}$ depends only on r modulo 2. In fact, we will find that

$$\chi_{g,r}^{(\ell)} = \begin{cases} \chi_g^{(\ell)}, & \text{for } r \equiv 0 \pmod 2; \\ \bar{\chi}_g^{(\ell)}, & \text{for } r \equiv 1 \pmod 2; \end{cases} \tag{4.3}$$

where $\chi_g^{(\ell)}$ and $\bar{\chi}_g^{(\ell)}$ are as defined in §3.4 and as tabulated in §B.2.

Group theoretically the multiplicities $\bar{\chi}_g^{(\ell)}$ and $\chi_g^{(\ell)}$ appearing in the shadow of $H_g^{(\ell)}$ are determined by the number of fixed points and anti-fixed points in the signed permutation representation of the group $G^{(\ell)}$, as explained in §3.2. For instance, we have $\bar{\chi}_g^{(\ell)} = \chi_g^{(\ell)} = \chi^{(\ell)}$ for the identity element $g = e$. From this interpretation we can deduce that the vanishing of $\bar{\chi}_g^{(\ell)}$ implies the vanishing of $\chi_g^{(\ell)}$, while it is possible to have $\chi_g^{(\ell)} = 0$ and $\bar{\chi}_g^{(\ell)} \neq 0$. It will also turn out that

$$\bar{\chi}_g^{(\ell)} = \chi_g^{(\ell)} \quad \text{unless} \quad 2|N_g. \tag{4.4}$$

For later use we define the combinations

$$\chi_{g,+}^{(\ell)} = \tfrac{1}{2}\big(\bar{\chi}_g^{(\ell)} + \chi_g^{(\ell)}\big), \quad \chi_{g,-}^{(\ell)} = \tfrac{1}{2}\big(\bar{\chi}_g^{(\ell)} - \chi_g^{(\ell)}\big), \tag{4.5}$$

which enumerate the number of fixed and anti-fixed points, respecitlvely, in the signed permutation representation of $G^{(\ell)}$ (cf. (3.3-3.4)). Of course in the cases where $\bar{\chi}_g^{(\ell)} = \chi_g^{(\ell)} = 0$ the function $H_g^{(\ell)} = \big(H_{g,r}^{(\ell)}\big)$ has vanishing shadow and is a vector-valued modular form in the usual sense.

Interestingly, just as in the $SL_2(\mathbb{Z})$ case that was considered in §2.3, the higher level vector-valued mock modular forms with shadows as described above are again closely related to the finite parts of meromorphic Jacobi forms of weight 1. More explicitly, from the transformation (2.29) and the simple fact $\theta_r^{(\ell)}(\tau, z + 1/2) = (-1)^r \theta_r^{(\ell)}(\tau, z)$ it is not difficult to check that the function

$$\psi_g^{(\ell)}(\tau, z) = \chi_{g,+}^{(\ell)} \mu_0^{(\ell)}(\tau, z) - \chi_{g,-}^{(\ell)} \mu_0^{(\ell)}(\tau, z + 1/2) + \sum_{r=1}^{\ell-1} H_{g,r}^{(\ell)}(\tau) \hat{\theta}_r^{(\ell)}(\tau, z) \tag{4.6}$$

transforms as a Jacobi form of weight 1 and index ℓ under the congruence subgroup $\Gamma_0(N_g)$ when $H_g^{(\ell)}$ is a mock modular form of weight 1/2 for $\Gamma_0(N_g)$ with shadow $S_g^{(\ell)} = \big(\chi_{g,r}^{(\ell)} S_r^{(\ell)}\big)$. Observe that, for those g with both $\chi_{g,+}^{(\ell)}$ and $\chi_{g,-}^{(\ell)}$ being non-zero, the weak Jacobi form $\psi_g^{(\ell)}(\tau, z)$ has a pole not only at $z=0$ but also at $z=1/2$. From

$$\mu_0^{(\ell)}(\tau,z)=\mathrm{Av}^{(\ell)}\left[\frac{y+1}{y-1}\right],\quad -\mu_0^{(\ell)}(\tau,z+1/2)=\mathrm{Av}^{(\ell)}\left[\frac{1-y}{1+y}\right],\tag{4.7}$$

we see that the last two terms of the decomposition given in (4.6) have the interpretation as the polar parts at the poles $z=0$ and $z=1/2$, respectively. In other words, we again have a decomposition $\psi_g^{(\ell)}=\psi_g^{(\ell),P}+\psi_g^{(\ell),F}$ into polar and finite parts given by

$$\begin{aligned}\psi_g^{(\ell),P}(\tau,z)&=\chi_{g,+}^{(\ell)}\mu_0^{(\ell)}(\tau,z)-\chi_{g,-}^{(\ell)}\mu_0^{(\ell)}(\tau,z+1/2),\\ \psi_g^{(\ell),F}(\tau,z)&=\sum_{r=1}^{\ell-1}H_{g,r}^{(\ell)}(\tau)\hat\theta_r^{(\ell)}(\tau,z),\end{aligned}\tag{4.8}$$

and the components $H_{g,r}^{(\ell)}$ of the mock modular form $H_g^{(\ell)}$ may again be interpreted as the theta-coefficients of a meromorphic Jacobi form; namely, $\psi_g^{(\ell)}(\tau,z)$.

Moreover, analogous to the $SL_2(\mathbb{Z})$ case, the mock modular form $H_g^{(\ell)}$ also enjoys a close relationship with a weight 0 index $\ell-1$ weak Jacobi form $Z_g^{(\ell)}$ which admits a decomposition into characters of the $N=4$ superconformal algebra at level $\ell-1$. To see this, observe that

$$\psi_g^{(\ell)}(\tau,z)=\frac{\chi_{g,+}^{(\ell)}}{\bar\chi_g^{(\ell)}}\,\Psi_{1,1}(\tau,z)\,Z_g^{(\ell)}(\tau,z)-\frac{\chi_{g,-}^{(\ell)}}{\bar\chi_g^{(\ell)}}\,\Psi_{1,1}(\tau,z+1/2)\,Z_g^{(\ell)}(\tau,z+1/2)\tag{4.9}$$

when $\bar\chi_g^{(\ell)}\neq 0$ and

$$\psi_g^{(\ell)}(\tau,z)=\Psi_{1,1}(\tau,z)\,Z_g^{(\ell)}(\tau,z)\tag{4.10}$$

when $\bar\chi_g^{(\ell)}=0$, where $\Psi_{1,1}(\tau,z)$ is as in (2.20) and $Z_g^{(\ell)}$ is the weak Jacobi form of weight 0 and index $\ell-1$ given by

$$Z_g^{(\ell)}(\tau,z)=\frac{1}{\Psi_{1,1}(\tau,z)}\left(\bar\chi_g^{(\ell)}\mu_0^{(\ell)}(\tau,z)+\sum_{0<r<\ell}\left(1+\frac{\bar\chi_g^{(\ell)}-\chi_{g,r}^{(\ell)}}{\chi_g^{(\ell)}}\right)H_{g,r}^{(\ell)}(\tau)\hat\theta_r^{(\ell)}(\tau,z)\right)\tag{4.11}$$

for $\chi_g^{(\ell)}\neq 0$ and

$$Z_g^{(\ell)}(\tau,z)=\frac{1}{\Psi_{1,1}(\tau,z)}\left(\bar\chi_g^{(\ell)}\mu_0^{(\ell)}(\tau,z)+\sum_{0<r<\ell}H_{g,r}^{(\ell)}(\tau)\hat\theta_r^{(\ell)}(\tau,z)\right)\tag{4.12}$$

otherwise.

Anticipating their relation to the conjugacy classes $[g]$ of $G^{(\ell)}$, another comment on the weight 0 forms $Z_g^{(\ell)}$ is in order here. As discussed in §3.2, if $\ell\in\Lambda$ and $\ell>2$ then $G^{(\ell)}$ has a unique central element z of order 2 and for any $g \in G^{(\ell)}$ the conjugacy casses $[g]$ and $[zg]$ are said to be paired, and a class $[g]$ is said to be self-paired if $[g] = [zg]$. We have

$$\bar{\chi}_g^{(\ell)} = \bar{\chi}_{zg}^{(\ell)}, \quad \chi_g^{(\ell)} = -\chi_{zg}^{(\ell)}, \tag{4.13}$$

and as a consequence $Z_g^{(\ell)}(\tau, z) = Z_{zg}^{(\ell)}(\tau, z)$. Therefore, while the vector-valued mock modular forms $H_g^{(\ell)}$ and the weight 1 meromorphic Jacobi forms $\psi_g^{(\ell)}(\tau, z)$ are generally distinct for different conjugacy classes $[g]$ of the group $G^{(\ell)}$, the weight 0 index $\ell - 1$ weak Jacobi forms $Z_g^{(\ell)}$ cannot distinguish between two classes that are paired in the above sense and are more naturally associated to the group $\bar{G}^{(\ell)}$ which is the quotient of $G^{(\ell)}$ by the subgroup $\langle z \rangle$.

Equipped with the $H^{(\ell)}$ defined in terms of decompositions of Jacobi forms as discussed in §2.5 it will develop that a convenient way to specify the vector-valued mock modular forms $H_g^{(\ell)} = \left(H_{g,r}^{(\ell)}\right)$ corresponding to non-identity conjugacy classes of $G^{(\ell)}$ will be to specify certain sets of weight 2 modular forms. To see how a weight 2 modular form is naturally associated with a vector-valued modular form with the properties described above, observe that we can eliminate the presence of the polar part in the weight 1 Jacobi form $\psi_g^{(\ell)}(\tau, z)$ (cf. (4.6)) by taking a linear combination with $\psi^{(\ell)}(\tau, z)$. To be more precise, note that

$$\hat{\psi}_g^{(\ell)}(\tau, z) = \sum_{r=1}^{\ell-1} \hat{H}_{g,r}^{(\ell)}(\tau) \hat{\theta}_r^{(\ell)}(\tau, z) = \psi_g^{(\ell)}(\tau, z) - \frac{\chi_{g,+}^{(\ell)}}{\chi^{(\ell)}} \psi^{(\ell)}(\tau, z) + \frac{\chi_{g,-}^{(\ell)}}{\chi^{(\ell)}} \psi^{(\ell)}(\tau, z + 1/2) \tag{4.14}$$

is a weight 1 index ℓ Jacobi form for $\Gamma_0(N_g)$ with no poles in z, and the functions

$$\hat{H}_{g,r}^{(\ell)}(\tau) = H_{g,r}^{(\ell)}(\tau) - \frac{\chi_{g,r}^{(\ell)}}{\chi^{(\ell)}} H_r^{(\ell)}(\tau) \tag{4.15}$$

are the components of a vector-valued modular form $\hat{H}_g^{(\ell)}$ for $\Gamma_0(N_g)$ in the usual sense (i.e. a mock modular form with vanishing shadow). From this we readily conclude that the function

$$F_g^{(\ell)}(\tau) = \sum_{r=1}^{\ell-1} \hat{H}_{g,r}^{(\ell)}(\tau) S_r^{(\ell)}(\tau) = -\frac{1}{4\pi i} \frac{\partial}{\partial z} \hat{\psi}_g^{(\ell)}(\tau, z)\big|_{z=0} \tag{4.16}$$

is a weight 2 modular form for the group $\Gamma_0(N_g)$.

For general values of ℓ, specifying the weight 2 form $F_g^{(\ell)}(\tau)$ is not sufficient to specify the whole vector-valued modular form $\hat{H}_g^{(\ell)}$ since we have evidently collapsed information in taking the particular combination (4.16). However, for $\ell \in \{2, 3\}$ we are in the privileged situation that specifying the weight 2 form $F_g^{(\ell)}(\tau)$ completely specifies all the components of $H_g^{(\ell)}$, as will be explained in more detail in the following section. For $\ell > 3$ we need more weight 2 forms, and we will also consider

$$F_g^{(\ell),2}(\tau) = \sum_{r=1}^{\ell-1} (-1)^{r+1} \hat{H}_{g,r}^{(\ell)}(\tau) S_{\ell-r}^{(\ell)}(\tau). \tag{4.17}$$

In the next section we give our concrete proposals for the McKay–Thompson series $H_g^{(\ell)}$ for all but a few of the conjugacy classes $[g]$ arising. We give closed expressions for all the $H_g^{(\ell)}$ in case $\ell \in \{2,3,4,5\}$ and we partially determine the $H_g^{(\ell)}$ for $\ell = \{7,13\}$. Although we do not offer analytic expressions for all the $H_g^{(\ell)}$ with $\ell = 7$ or $\ell = 13$ we have predictions for the low degree terms in the Fourier developments of all the McKay–Thompson series at all lambencies $\ell \in \Lambda$ and these are detailed in the tables of §C.

4.2 Lambency Two

$[g]$	N_g	$F_g^{(2)}(\tau)$
$1A$	1	0
$2A$	2	$-16\Lambda_2$
$2B$	4	$24\Lambda_2 - 8\Lambda_4 = -2\eta(\tau)^8/\eta(2\tau)^4$
$3A$	3	$-6\Lambda_3$
$3B$	9	$-2\eta(\tau)^6/\eta(3\tau)^2$
$4A$	8	$-4\Lambda_2 + 6\Lambda_4 - 2\Lambda_8 = -2\eta(2\tau)^8/\eta(4\tau)^4$
$4B$	4	$4(\Lambda_2 - \Lambda_4)$
$4C$	16	$-2\eta(\tau)^4\eta(2\tau)^2/\eta(4\tau)^2$
$5A$	5	$-2\Lambda_5$
$6A$	6	$2(\Lambda_2 + \Lambda_3 - \Lambda_6)$
$6B$	36	$-2\eta(\tau)^2\eta(2\tau)^2\eta(3\tau)^2/\eta(6\tau)^2$
$7AB$	7	$-\Lambda_7$
$8A$	8	$\Lambda_4 - \Lambda_8$
$10A$	20	$-2\eta(\tau)^3\eta(2\tau)\eta(5\tau)/\eta(10\tau)$
$11A$	11	$\frac{2}{5}(-\Lambda_{11} + 11f_{11})$
$12A$	24	$-2\eta(\tau)^3\eta(4\tau)^2\eta(6\tau)^3/\eta(2\tau)\eta(3\tau)\eta(12\tau)^2$
$12B$	144	$-2\eta(\tau)^4\eta(4\tau)\eta(6\tau)/\eta(2\tau)\eta(12\tau)$
$14AB$	14	$\frac{1}{3}(\Lambda_2 + \Lambda_7 - \Lambda_{14} + 14f_{14})$
$15AB$	15	$\frac{1}{4}(\Lambda_3 + \Lambda_5 - \Lambda_{15} + 15f_{15})$
$21AB$	63	$\frac{1}{3}(-7\eta(\tau)^3\eta(7\tau)^3/\eta(3\tau)\eta(21\tau) + \eta(\tau)^6/\eta(3\tau)^2)$
$23AB$	23	$\frac{1}{11}(-\Lambda_{23} + 23f_{23,a} + 69f_{23,b})$

Table 3: The list of weight 2 modular forms $F_g^{(2)}(\tau)$ for $\Gamma_0(N_g)$.

When $\ell = 2$ the vector-valued mock modular form $H_g^{(2)}$ has only one component $H_{g,1}^{(2)}$ and our conjecture relating the $H_g^{(2)}$ and the group $G^{(2)}$ is nothing but the conjecture relating (scalar-valued) mock modular forms and the largest Mathieu group M_{24} that has been investigated recently [15, 19, 24, 20, 21, 22, 108, 109]. See also [41, 110] for reviews and [111, 112] for related discussions. Explicit expressions for the McKay–Thompson series arising from the M_{24}-module that is conjectured to underlie this connection have been proposed in [19, 20, 21, 22]. As mentioned before, one convenient way to express them is via a set of weight 2 modular forms $F_g^{(2)}(\tau)$. In this case (4.15-4.16) simply reduces to

$$H_{g,1}^{(2)}(\tau) = \frac{\chi_g^{(2)}}{24} H_1^{(2)}(\tau) + \frac{F_g^{(2)}(\tau)}{\eta(\tau)^3} \ , \tag{4.18}$$

where we have used $\chi^{(2)} = 24$ and $S_1^{(2)}(\tau) = \eta(\tau)^3$. For later use and for the sake of completeness we collect the explicit expressions for $F_g^{(2)}(\tau)$ for all conjugacy classes $[g] \subset M_{24}$ in Table 3. They are given in terms of eta quotients and standard generators of the weight 2 modular forms of level N. Among the latter are those denoted here by $\Lambda_N(\tau)$ and $f_N(\tau)$ and given explicitly in Appendix A.

4.3 Lambency Three

We would like to specify the two components $H_{g,1}^{(3)}(\tau)$ and $H_{g,2}^{(3)}(\tau)$ of the vector-valued mock modular form $\big(H_{g,r}^{(3)}(\tau)\big)$ which we propose to be the McKay–Thompson series arising from the $G^{(3)}$-module $K^{(3)}$. As mentioned earlier, to specify $H_{g,1}^{(3)}(\tau)$ and $H_{g,2}^{(3)}(\tau)$ it is sufficient to specify the weight 2 forms defined in (4.16). To see this, recall that to any given conjugacy class $[g]$ we can associate a conjugacy class $[zg]$ such that (4.13) holds. From this we obtain

$$\begin{aligned}
H_{g,1}^{(3)}(\tau) &= \frac{\bar{\chi}_g^{(3)}}{\chi^{(3)}} H_1^{(3)}(\tau) + \frac{1}{2}\frac{\eta(4\tau)^2}{\eta(2\tau)^5}\big(F_g^{(3)}(\tau) + F_{zg}^{(3)}(\tau)\big),\\
H_{g,2}^{(3)}(\tau) &= \frac{\chi_g^{(3)}}{\chi^{(3)}} H_2^{(3)}(\tau) + \frac{1}{4}\frac{\eta(2\tau)}{\eta(\tau)^2\eta(4\tau)^2}\big(F_g^{(3)}(\tau) - F_{zg}^{(3)}(\tau)\big),
\end{aligned} \tag{4.19}$$

where we have used the eta quotient expressions

$$S_1^{(3)}(\tau) = \frac{\eta(2\tau)^5}{\eta(4\tau)^2}, \quad S_2^{(3)}(\tau) = 2\frac{\eta(\tau)^2\eta(4\tau)^2}{\eta(2\tau)}, \tag{4.20}$$

for the components of the unary theta series at $\ell = 3$.

The explicit expressions for $F_g^{(3)}(\tau)$ are listed in Table 4.

We also note that for the classes $3B$ and $6B$ we also have the following alternative expressions

$[g]$	N_g	$F_g^{(3)}(\tau)$
1A	1	0
2A	4	0
4A	16	$-2\eta(\tau)^4\eta(2\tau)^2/\eta(4\tau)^2$
2B	2	$-16\Lambda_2$
2C	4	$16(\Lambda_2 - \Lambda_4/3)$
3A	3	$-6\Lambda_3$
6A	12	$-9\Lambda_2 - 2\Lambda_3 + 3\Lambda_4 + 3\Lambda_6 - \Lambda_{12}$
3B	9	$8\Lambda_3 - 2\Lambda_9 + 2\,\eta^6(\tau)/\eta^2(3\tau)$
6B	36	$-2\eta(\tau)^5\eta(3\tau)/\eta(2\tau)\eta(6\tau)$
4B	8	$-2\eta(2\tau)^8/\eta(4\tau)^4$
4C	4	$-8\Lambda_4/3$
5A	5	$-2\Lambda_5$
10A	20	$\sum_{d\vert 20} c_{10A}(d)\Lambda_d + \frac{20}{3}f_{20}$
12A	144	$-2\eta(\tau)\eta(2\tau)^5\eta(3\tau)/\eta(4\tau)^2\eta(6\tau)$
6C	6	$2(\Lambda_2 + \Lambda_3 - \Lambda_6)$
6D	12	$-5\Lambda_2 - 2\Lambda_3 + \frac{5}{3}\Lambda_4 + 3\Lambda_6 - \Lambda_{12}$
8AB	32	$-2\eta(2\tau)^4\eta(4\tau)^2/\eta(8\tau)^2$
8CD	8	$-2\Lambda_2 + \frac{5}{3}\Lambda_4 - \Lambda_8$
20AB	80	$-2\eta(2\tau)^7\eta(5\tau)/\eta(\tau)\eta(4\tau)^2\eta(10\tau)$
11AB	11	$-\frac{2}{5}\Lambda_{11} - \frac{33}{5}f_{11}$
22AB	44	$\sum_{d\vert 44} c_{22AB}(d)\Lambda_d(\tau) - \frac{11}{5}\sum_{d\vert 4} c'_{22AB}(d)f_{11}(d\tau) + \frac{22}{3}f_{44}(\tau)$

$c_{10A}(d) = -5, \frac{5}{3}, -\frac{2}{3}, 1, -\frac{1}{3}$ for $d = 2, 4, 5, 10, 20$
$c_{22AB}(d) = -\frac{11}{5}, \frac{11}{15}, -\frac{2}{15}, \frac{1}{5}, -\frac{1}{15}$ for $d = 2, 4, 11, 22, 44$
$c'_{22AB}(d) = 1, 4, 8$ for $d = 1, 2, 4$

Table 4: The list of weight 2 modular forms $F_g^{(3)}(\tau)$ for $\Gamma_0(N_g)$.

for the McKay–Thompson series in terms of eta quotients:

$$H^{(3)}_{3B,1}(\tau) = H^{(3)}_{6B,1}(\tau) = -2\frac{\eta(\tau)\eta(6\tau)^5}{\eta(3\tau)^3\eta(12\tau)^2}, \quad H^{(3)}_{3B,2}(\tau) = -H^{(3)}_{6B,2}(\tau) = -4\frac{\eta(\tau)\eta(12\tau)^2}{\eta(3\tau)\eta(6\tau)}. \quad (4.21)$$

Coincidences with Ramanujan's mock theta functions will be discussed (4.7).

4.4 Lambency Four

The McKay–Thompson series for lambency 4 display an interesting relation with those for lambency 2. To see this, notice the following relation among the theta functions

$$S^{(4)}_1(2\tau) - S^{(4)}_3(2\tau) = S^{(2)}_1(\tau). \quad (4.22)$$

Given this relation it is natural to consider the function

$$H^{(4)}_{g,*}(\tau) := H^{(4)}_{g,1}(\tau) - H^{(4)}_{g,3}(\tau) \quad (4.23)$$

for each conjugacy classe $[g]$ of $G^{(4)}$. Note that $H^{(4)}_{g,1}(\tau)$ and $H^{(4)}_{g,3}(\tau)$ can be reconstructed from $H^{(4)}_{g,*}(\tau)$ since they are q-series of the form $q^{-1/16}$ times a series of even or odd powers of $q^{1/2}$, respectively. Explicitly, we have

$$\begin{aligned} H^{(4)}_{g,1}(\tau) &= \tfrac{1}{2}\big(H^{(4)}_{g,*}(\tau) + e(\tfrac{1}{16})H^{(4)}_{g,*}(\tau+1)\big),\\ H^{(4)}_{g,3}(\tau) &= \tfrac{1}{2}\big(-H^{(4)}_{g,*}(\tau) + e(\tfrac{1}{16})H^{(4)}_{g,*}(\tau+1)\big). \end{aligned} \quad (4.24)$$

In order to obtain an expression for $H^{(4)}_{g,*}(\tau)$ we rely on the following two observations. First, recall in §3.4 we have observed a relation between the Frame shapes of $g \in G^{(4)}$ and $g' \in G^{(2)}$. It turns out that for a pair of group elements $g \in G^{(4)}$ and $g' \in G^{(2)}$ related in the way described in Proposition 3.4 their McKay–Thompson series $H^{(4)}_{g}(2\tau)$ and $H^{(2)}_{g'}(\tau)$ are also related in a simple way. As examples of this we have

$$\begin{aligned}
H^{(4)}_{1A,*}(\tau) &= H^{(4)}_{2A,*}(\tau) = H^{(2)}_{2A}(\tau/2)\\
H^{(4)}_{2B,*}(\tau) &= H^{(2)}_{4A}(\tau/2)\\
H^{(4)}_{2C,*}(\tau) &= H^{(2)}_{4B}(\tau/2)\\
H^{(4)}_{3A,*}(\tau) &= H^{(4)}_{6A,*}(\tau) = H^{(2)}_{6A}(\tau/2)\\
H^{(4)}_{4C,*}(\tau) &= H^{(2)}_{8A}(\tau/2)\\
H^{(4)}_{6BC,*}(\tau) &= H^{(2)}_{12A}(\tau/2)\\
H^{(4)}_{7AB,*}(\tau) &= H^{(4)}_{14AB,*}(\tau) = H^{(2)}_{14AB}(\tau/2).
\end{aligned} \tag{4.25}$$

This leaves us just the classes $4A$, $4B$, and $8A$ that are not related to any element of $G^{(2)} \simeq M_{24}$ in the way described in Proposition 3.4. All of these classes have $\chi^{(4)}_g = \bar{\chi}^{(4)}_g = 0$. A second observation is that all the $H^{(3)}_{g,r}(\tau)$ for those classes $[g] \subset G^{(3)}$ with $\chi^{(3)}_g = \bar{\chi}^{(3)}_g = 0$ have an expression in terms of eta quotients according to Table 4, and so do the $H^{(2)}_g(\tau)$ for those classes $[g] \subset G^{(2)}$ with $\chi^{(2)}_g = 0$. This is consistent with our expectation that the shadow of the vector-valued mock modular form $\big(H^{(\ell)}_{g,r}(\tau)\big)$ is given by $\big(\chi^{(\ell)}_{g,r} S^{(\ell)}_r(\tau)\big)$, and hence $\big(H^{(\ell)}_{g,r}(\tau)\big)$ is nothing but a vector-valued modular form in the usual sense when $\chi^{(\ell)}_g = \bar{\chi}^{(\ell)}_g = 0$. Given this it is natural to ask whether we can find similar eta quotient expressions for $H^{(4)}_{g,*}(\tau)$ with $\chi^{(\ell)}_g = \bar{\chi}^{(\ell)}_g = 0$ when $\ell > 3$. We find

$$\begin{aligned}
H^{(4)}_{4A,*}(\tau) &= -2\frac{\eta(\tau/2)\eta(\tau)^2}{\eta(2\tau)^2},\\
H^{(4)}_{4B,*}(\tau) &= -2\frac{\eta(\tau/2)\eta(\tau)^4}{\eta(\tau)^2\eta(4\tau)^2},\\
H^{(4)}_{8A,*}(\tau) &= -2\frac{\eta(\tau)^3}{\eta(\tau/2)\eta(4\tau)}.
\end{aligned} \tag{4.26}$$

Together with (4.24) and Table 3, the above equations completely specify $H^{(4)}_{g,1}(\tau)$ and $H^{(4)}_{g,3}(\tau)$ for all conjugacy classes $[g]$ of $G^{(4)}$. We are left to determine the second components $H^{(4)}_{g,2}(\tau)$. To start with, we have $H^{(4)}_2(\tau) = H^{(4)}_{1A,2}(\tau) = -H^{(4)}_{2A,2}(\tau)$ given in terms of the decomposition of the weight 0 Jacobi form as explained in §2.5. To determine the rest, notice that (4.13) and (4.16) implies that

$$S^{(4)}_2(\tau)\left(H^{(4)}_{g,2}(\tau) - \frac{\chi^{(4)}_{g,2}}{\chi^{(4)}}H^{(4)}_2(\tau)\right) \tag{4.27}$$

should be a weight 2 modular form. Employing this consideration we arrive at

$$\begin{aligned}
H^{(4)}_{3A,2}(\tau) &= -H^{(4)}_{6A,2}(\tau)\\
&= \frac{1}{4}H^{(4)}_2(\tau) + \frac{1}{2\eta(2\tau)^3}\Big(-3\Lambda_2(\tau) - 4\Lambda_3(\tau) + \Lambda_6(\tau)\Big),\\
H^{(4)}_{7AB,2}(\tau) &= -H^{(4)}_{14AB,2}(\tau)\\
&= \frac{1}{8}H^{(4)}_2(\tau) + \frac{1}{12\,\eta(2\tau)^3}\Big(-7\Lambda_2(\tau) - 4\Lambda_7(\tau) + \Lambda_{14}(\tau) + 28f_{14}(\tau)\Big),
\end{aligned} \tag{4.28}$$

where we have used $S^{(4)}_2(\tau) = 2\eta(2\tau)^3$ and the hypothesis that

$$H^{(4)}_{g,2}(\tau) = 0 \tag{4.29}$$

for all classes not in the set $\{1A, 2A, 3A, 6A, 7AB, 14AB\}$. This finishes our proposal for the McKay–Thompson series $\big(H^{(4)}_{g,r}\big)$ at $\ell = 4$. Coincidences between some of the $H^{(4)}_{g,r}$ and Ramanujan's mock theta functions will be discussed in §4.7.

4.5 Lambency Five

As mentioned before, for $\ell = 5$ it is no longer sufficient to specify the weight 2 forms $F^{(\ell)}_g(\tau)$ as defined in (4.16). In this case we also need to specify the weight 2 forms $F^{(\ell),2}_g(\tau)$ defined in (4.17).

With the data of $F^{(5)}_g(\tau)$ and $F^{(5),2}_g(\tau)$ we can determine the $H^{(5)}_{g,r}$ using the relations

$$H^{(5)}_{g,r}(\tau) = \hat{H}^{(5)}_{g,r}(\tau) + \frac{\chi^{(5)}_{g,r}}{6}H^{(5)}_r(\tau) \tag{4.30}$$

and

$$\begin{aligned}
\hat{H}^{(5)}_{g,1}(\tau)S^{(5)}_1(\tau) + \hat{H}^{(5)}_{g,3}(\tau)S^{(5)}_3(\tau) &= \tfrac{1}{2}\Big(F^{(5)}_g(\tau) + F^{(5)}_{zg}(\tau)\Big)\\
\hat{H}^{(5)}_{g,1}(\tau)S^{(5)}_4(\tau) + \hat{H}^{(5)}_{g,3}(\tau)S^{(5)}_2(\tau) &= \tfrac{1}{2}\Big(F^{(5),2}_g(\tau) + F^{(5),2}_{zg}(\tau)\Big)\\
\hat{H}^{(5)}_{g,2}(\tau)S^{(5)}_2(\tau) + \hat{H}^{(5)}_{g,4}(\tau)S^{(5)}_4(\tau) &= \tfrac{1}{2}\Big(F^{(5)}_g(\tau) - F^{(5)}_{zg}(\tau)\Big)\\
\hat{H}^{(5)}_{g,2}(\tau)S^{(5)}_3(\tau) + \hat{H}^{(5)}_{g,4}(\tau)S^{(5)}_1(\tau) &= \tfrac{1}{2}\Big(-F^{(5),2}_g(\tau) + F^{(5),2}_{zg}(\tau)\Big)
\end{aligned} \tag{4.31}$$

where g and zg again form a pair satisfying (4.13).

Moreover, notice that in $F^{(5),2}_g(\tau)$ the part of the sum involving even r has an expansion in $q^{1/4+k}$ for non-negative integers k, while the sum over odd r has an expansion in $q^{3/4+k}$. As

$[g]$	N_g	$F_g^{(5)}(\tau)$	$F_g^{(5),2}(\tau)$
1A	1	0	0
2A	4	0	0
2B	4	$16(\Lambda_2-\Lambda_4/3)$	$-\frac{8}{3}\eta(\tau)^8/\eta(\tau/2)^4$
2C	2	$-16\Lambda_2$	$e(\frac{1}{4})\,F_{2B}^{(5),2}(\tau+1)$
3A	9	$-2\eta^6(\tau)/\eta^2(3\tau)$	$-2\frac{\eta(\tau)^8\eta(3\tau/2)^2\eta(6\tau)^2}{\eta(\tau/2)^2\eta(2\tau)^2\eta(3\tau)^4}$
6A	36	$-2\,\eta(\tau)^2\eta(2\tau)^2\eta(3\tau)^2/\eta(6\tau)^2$	$2\,\eta(\tau/2)^2\eta(\tau)^2\eta(3\tau)^2/\eta(3\tau/2)^2$
4AB	16	$-2\eta(2\tau)^{14}/\eta(\tau)^4\eta(4\tau)^6$	$-16\eta(\tau)^2\eta(4\tau)^6/\eta(2\tau)^4$
4CD	8	$-8/3\Lambda_4(\tau)$	$-\frac{32}{3}\,\eta(2\tau)^2\eta(4\tau)^2/\eta(\tau)^2$
5A	5	$-2\Lambda_5$	$e(\frac{1}{4})\,F_{10A}^{(5),2}(\tau+1)$
10A	20	$\sum_{d\mid 20}c_{10A}(d)\Lambda_d-\frac{40}{3}f_{20}$	$\sum_{d\mid 20}c_{10A,2}(d)\Lambda_d(\tau/4)+\frac{10}{3}f_{20}(\tau/4)$
12AB	144	$-2\eta(\tau)^2\eta(2\tau)^2\eta(6\tau)^4/\eta(3\tau)^2\eta(12\tau)^2$	$-4\eta(\tau)^2\eta(2\tau)^2\eta(12\tau)^2/\eta(6\tau)^2$
$c_{10A}(d)=-5,\frac{5}{3},-\frac{2}{3},1,-\frac{1}{3}$ for $d=2,4,5,10,20$			
$c_{10A,2}(d)=-\frac{5}{4},\frac{5}{24},-\frac{1}{3},\frac{1}{4},-\frac{1}{24}$ for $d=2,4,5,10,20$			

Table 5: The list of weight 2 modular forms $F_g^{(5)}(\tau)$ on $\Gamma_0(N_g)$ and the other set of weight 2 modular forms $F_g^{(5),2}(\tau)$.

a result, it is sufficient to specify $F_g^{(\ell),2}(\tau)$ for one of the paired classes g and zg as they are related to each other by

$$F_{zg}^{(5),2}(\tau)=e(\tfrac{1}{4})\,F_g^{(\ell),2}(\tau+1)\ . \tag{4.32}$$

Using (4.30-4.32), the data recorded in Table 5 are sufficient to explicitly specify all the mock modular forms $H_g^{(5)}(\tau)$ for all $[g]\subset G^{(5)}$ at lambency 5.

4.6 Lambencies Seven and Thirteen

For $\ell=7$, apart from the $1A$ and $2A$ classes whose McKay-Thomson series have been given in terms of the weight 0 Jacobi form $Z^{(7)}$ in §2.5, there are five more classes $3AB$, $4A$, and $6AB$ whose McKay-Thomson series $H_g^{(7)}$ we would like to identify. We specify the weight 2 forms associated to these classes. Notice that these expressions are not sufficient to completely determine all the components $H_{g,r}^{(7)}$. Nevertheless, they provide strong evidence for the mock modular properties of the proposed McKay–Thompson series.

$$\begin{aligned}
F^{(7)}_{4A}(\tau) &= -2\frac{\eta(\tau)^4\eta(2\tau)^2}{\eta(4\tau)^2},\\
F^{(7),2}_{4A}(\tau) &= 4\frac{\eta(\tau)^6\eta(4\tau)^2}{\eta(2\tau)^4},\\
F^{(7)}_{3AB}(\tau) &= -6\Lambda_3(\tau),\\
F^{(7)}_{6AB}(\tau) &= -9\Lambda_2(\tau)-2\Lambda_3(\tau)+3\Lambda_4(\tau)+3\Lambda_6(\tau)-\Lambda_{12}(\tau),\\
F^{(7),2}_{6AB}(4\tau) &= -\frac{9}{4}\Lambda_2(\tau)-\Lambda_3(\tau)+\frac{3}{8}\Lambda_4(\tau)+\frac{3}{4}\Lambda_6(\tau)-\frac{1}{8}\Lambda_{12}(\tau).
\end{aligned}\tag{4.33}$$

Similarly, for lambency 13 we have

$$\begin{aligned}
F^{(13)}_{4AB}(\tau) &= -2\frac{\eta(2\tau)^{14}}{\eta(\tau)^4\eta(4\tau)^6},\\
F^{(13),2}_{4AB}(\tau) &= -16\frac{\eta(\tau)^2\eta(4\tau)^6}{\eta(2\tau)^4}.
\end{aligned}\tag{4.34}$$

4.7 Mock Theta Functions

As mentioned in §2.1 we observe that in many cases across different lambencies the components $H^{(\ell)}_{g,r}$ of the McKay–Thompson series coincide with known mock theta functions. In particular, we often encounter Ramanujan's mock theta functions, and always in such a way that the order is divisible by the lambency. In this section we will list these conjectural identities between $H^{(\ell)}_{g,r}$ and mock theta functions identified previously in the literature.

For $\ell = 2$ two of the functions $H^{(2)}_g(\tau)$ are related to Ramanujan's mock theta functions of orders 2 and 8 (cf. (2.7)) through

$$\begin{aligned}
H^{(2)}_{4B}(\tau) &= -2q^{-1/8}\mu(q),\\
H^{(2)}_{8A}(\tau) &= -2q^{-1/8}U_0(q).
\end{aligned}\tag{4.35}$$

For $\ell = 3$ we encounter the following order 3 mock theta functions of Ramanujan:

$$\begin{aligned}
H^{(3)}_{2B,1}(\tau) = H^{(3)}_{2C,1}(\tau) = H^{(3)}_{4C,1}(\tau) &= -2q^{-1/12}f(q^2),\\
H^{(3)}_{6C,1}(\tau) = H^{(3)}_{6D,1}(\tau) &= -2q^{-1/12}\chi(q^2),\\
H^{(3)}_{8C,1}(\tau) = H^{(3)}_{8D,1}(\tau) &= -2q^{-1/12}\phi(-q^2),\\
H^{(3)}_{2B,2}(\tau) = -H^{(3)}_{2C,2}(\tau) &= -4q^{2/3}\omega(-q),\\
H^{(3)}_{6C,2}(\tau) = -H^{(3)}_{6D,2}(\tau) &= 2q^{2/3}\rho(-q).
\end{aligned}\tag{4.36}$$

(Cf. (2.9).) The description of the shadow of $H_g^{(3)}$ is consistent with the fact that, among the seven order 3 mock theta functions of Ramanujan, $f(\tau)$, $\phi(\tau)$, $\psi(\tau)$ and $\chi(\tau)$ form a group with the same shadow $(S_1^{(3)}(\tau))$ while the other three $\omega(\tau)$, $\nu(\tau)$ and $\rho(\tau)$ form another group with another shadow $(S_2^{(3)}(\tau))$. Moreover, various relations among these order 3 mock theta functions can be obtained as consequences of the above identification.

For $\ell = 4$ we encounter the following order 8 mock theta functions

$$\begin{aligned} H^{(4)}_{2C,1}(\tau) &= q^{-\frac{1}{16}}\big(-2S_0(q)+4T_0(q)\big),\\ H^{(4)}_{2C,3}(\tau) &= q^{\frac{7}{16}}\big(2S_1(q)-4T_1(q)\big),\\ H^{(4)}_{4C,1}(\tau) &= -2\,q^{-\frac{1}{16}}S_0(q),\\ H^{(4)}_{4C,3}(\tau) &= 2\,q^{\frac{7}{16}}\,S_1(q). \end{aligned} \tag{4.37}$$

(Cf. (2.11).) Comparing with (2.7) we see that our proposal implies the identities

$$\begin{aligned} \mu(q) &= S_0(q^2)-2T_0(q^2)+q\big(S_1(q^2)-2T_1(q^2)\big)\\ &= U_0(q)-2U_1(q),\\ U_0(q) &= S_0(q^2)+q\,S_1(q^2),\\ U_1(\tau) &= T_0(q^2)+q\,T_1(q^2), \end{aligned} \tag{4.38}$$

between different mock theta functions of Ramanujan. See, for instance, [113] for a collection of such identities.

For $\ell = 5$ we encounter four of Ramanujan's order 10 mock theta functions:

$$\begin{aligned} H^{(5)}_{2BC,1}(\tau) &= H^{(5)}_{4CD,1}(\tau) = -2q^{-\frac{1}{20}}\,X(q^2),\\ H^{(5)}_{2BC,3}(\tau) &= H^{(5)}_{4CD,3}(\tau) = -2q^{-\frac{9}{20}}\,\chi_{10}(q^2),\\ H^{(5)}_{2C,2}(\tau) &= -H^{(5)}_{2B,2}(\tau) = 2q^{-\frac{1}{5}}\,\psi_{10}(-q),\\ H^{(5)}_{2C,4}(\tau) &= -H^{(5)}_{2B,4}(\tau) = -2q^{\frac{1}{5}}\,\phi_{10}(-q). \end{aligned} \tag{4.39}$$

(Cf. (2.10).)

4.8 Automorphy

In this section we discuss the automorphy of the proposed McKay–Thompson series $H_g^{(\ell)} = \big(H_{g,r}^{(\ell)}\big)$. As mentioned in §4.1, the function $H_g^{(\ell)}$ is a vector-valued mock modular form with shadow $\big(\chi_{g,r}S_r^{(\ell)}\big)$ for some $\Gamma_g \subset SL_2(\mathbb{Z})$ with a certain (matrix-valued) multiplier ν_g. In this subsection we will specify the group Γ_g and the multiplier ν_g. The multipliers we specify here can be verified explicitly using the data given in §§4.2-4.6 (except for the few conjugacy classes at $\ell = 7$ and $\ell = 13$ for which the McKay–Thompson series $H_g^{(\ell)}$ have not been completely determined).

We find that the automorphy of the vector-valued function $H_g^{(\ell)}$ is governed, in way that we shall describe presently, by the signed permutation representation of $G^{(\ell)}$ arising from the construction (as a subgroup of Oct_m for $m = 24/(\ell-1)$) given in §3.4. We use this representation to define the symbols $n_g|h_g$ which appear in the second row of each table 20-47, and also in the twisted Euler character tables 14-19, and we will explain below how to use these symbols to determine the multiplier system for each $H_g^{(\ell)}$. It will develop also that the twisted Euler character $g \mapsto \chi_g^{(\ell)}$ (cf. Tables 14-19) attached to the signed permutation representation of $G^{(\ell)}$ determines the shadow of $H_g^{(\ell)}$.

Recall from §3.4 that the signed permutation representation of $G^{(\ell)}$ naturally induces a permutation representation of the same degree, and this permutation representation factors through the quotient $\bar{G}^{(\ell)} = G^{(\ell)}/2$ of $G^{(\ell)}$ by its unique central subgroup of order 2 (except in case $\ell = 2$ when the signed and unsigned permutation representations coincide, and we have $G^{(2)} = \bar{G}^{(2)}$). Let $g \mapsto \bar{g}$ denote the natural map $G^{(\ell)} \to \bar{G}^{(\ell)}$ and observe that (for $\ell > 2$) if the unique central involution of $G^{(\ell)}$ belongs to the cyclic subgroup $\langle g \rangle$ generated by g then $o(g) = 2o(\bar{g})$ and otherwise $o(g) = o(\bar{g})$. We say that $g \in G^{(\ell)}$ is *split* over $\bar{G}^{(\ell)}$ in case $o(g) = o(\bar{g})$ and we call g *non-split* otherwise. (By this definition every element of $G^{(2)} = \bar{G}^{(2)}$ is split.)

Recall from §3.4 that $g \mapsto \Pi_g^{(\ell)}$ denotes the map attaching signed permutation Frame shapes to elements of $G^{(\ell)}$ and $g \mapsto \bar{\Pi}_g^{(\ell)}$ denotes the Frame shapes (actually cycle shapes) arising from the (unsigned) permutation representation (on $m = 24/(\ell-1)$ points). Taking a formal product of Frame shapes $\tilde{\Pi}_g^{(\ell)} = \Pi_g^{(\ell)} \bar{\Pi}_g^{(\ell)}$ (defined so that $j^{m_1} j^{m_2} = j^{m_1+m_2}$, &c.) we obtain the Frame shape of $g \in G^{(\ell)}$ regarded as a permutation of the $2m$ points $\{\pm e_i\}$ for $i \in \Omega^{(\ell)}$ (cf. §3.4). In particular, $\tilde{\Pi}_g$ is a cycle shape and none of the exponents appearing in $\tilde{\Pi}_g$ are negative. Given $g \in G^{(\ell)}$ and $\tilde{\Pi}_g = j_1^{m_1} \cdots j_k^{m_l}$ with $j_1 < \cdots < j_k$ and $m_i > 0$ define $N_g = j_1 j_k$. That is, set N_g to be the product of the shortest and longest cycle lengths appearing in a cycle decomposition for g regarded as a permutation on the $2m$ points $\{\pm e_i\}$. Now define the symbols $n_g|h_g$ by setting $n_g = o(\bar{g})$ for all $g \in G^{(\ell)}$ and all $\ell \in \Lambda$, and by setting $h_g = N_g/n_g$ for all $g \in G^{(\ell)}$ and $\ell \in \Lambda$ except when $\ell = 4$ and g is non-split in which case set $h_g = N_g/2n_g$. The symbols $n_g|h_g$ are specified in Tables 14-19, and also in Tables 20-47, in the rows labelled Γ_g. We omit the $|h_g$ when $h_g = 1$, so n_g is a shorthand for $n_g|1$ in these tables.

The significance of the value n_g is that it is the minimal positive integer n for which $H_g^{(\ell)}$ is a mock modular form (of weight 1/2) on $\Gamma_0(n)$, and the significance of h_g is that, as we shall see momentarily, it is the minimal positive integer h for which the multiplier for $H_g^{(\ell)}$ coincides with the conjugate multiplier for the (vector-valued) cusp form $S^{(\ell)}$ when restricted to $\Gamma_0(nh)$ (for $n = n_g$). Since $n_g h_g \leq N_g$ for all g the multiplier for $H_g^{(\ell)}$ coincides with the conjugate multiplier for $S^{(\ell)}$ when regarded as a mock modular form on $\Gamma_0(N_g)$ for all g. It is very curious that this coincidence of multipliers extends to the larger group $\Gamma_0(N_g/2)$ in the case that $\ell = 4$ and g is non-split (i.e. $o(g) = 2o(\bar{g})$).

Table 6: Admissible $v^{(\ell)}$

ℓ	2	3	4	5	7	13
$v^{(\ell)}$	1	5	3	7	1	7

Given a pair of positive integers (n, h) we define a matrix-valued function $\rho^{(\ell)}_{n|h}$ on $\Gamma_0(n)$ as follows. For each $\ell \in \Lambda$ let the integer $v^{(\ell)}$ be as specified in Table 6. When h divides n we set

$$\rho^{(\ell)}_{n|h}(\gamma) = \mathrm{e}\left(-v^{(\ell)}\frac{cd}{nh}\right)\mathrm{I}_{\ell-1} \tag{4.40}$$

where $\mathrm{I}_{\ell-1}$ denotes the $(\ell-1)\times(\ell-1)$ identity matrix. When h does not divide n and n is even we set

$$\rho^{(\ell)}_{n|h}(\gamma) = \mathrm{e}\left(-v^{(\ell)}\frac{cd}{nh}\frac{(n,h)}{n}\right)\mathrm{J}_{\ell-1}^{c(d+1)/n}\mathrm{K}_{\ell-1}^{c/n} \tag{4.41}$$

where (n, h) denotes the greatest common divisor of n and h, and $\mathrm{J}_{\ell-1}$ and $\mathrm{K}_{\ell-1}$ are the $(\ell-1)\times(\ell-1)$ matrices given by

$$\mathrm{J}_{\ell-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (-1)^{\ell} \end{pmatrix}, \quad \mathrm{K}_{\ell-1} = \begin{pmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & .\cdot\cdot & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 & 0 \end{pmatrix}, \tag{4.42}$$

and when h does not divide n and n is odd we set

$$\rho^{(\ell)}_{n|h}(\gamma) = \mathrm{e}\left(-v^{(\ell)}\frac{cd}{nh}\frac{n}{(n,h)}\right)\mathrm{J}_{\ell-1}^{c(d+1)/n}\mathrm{K}_{\ell-1}^{c/n}. \tag{4.43}$$

Now the multiplier system $\nu_g^{(\ell)}$ for the umbral mock modular form $H_g^{(\ell)}$ is the matrix-valued function on $\Gamma_0(n) = \Gamma_0(n_g)$ given simply by

$$\nu_g^{(\ell)} = \nu^{(\ell)}_{n|h} = \rho^{(\ell)}_{n|h}\overline{\sigma^{(\ell)}} \tag{4.44}$$

where $n = n_g$ and $h = h_g$ and $\sigma^{(\ell)} = (\sigma^{(\ell)}_{ij})$ denotes the (matrix-valued) multiplier system for the (vector-valued) theta series $S^{(\ell)}$ (cf. (2.26)) and satisfies

$$\sigma^{(\ell)}(\gamma)S^{(\ell)}(\gamma\tau)\mathrm{jac}(\gamma,\tau)^{3/4} = S^{(\ell)}(\tau) \tag{4.45}$$

for $\gamma \in \Gamma_0(1) = SL_2(\mathbb{Z})$ where $\mathrm{jac}(\gamma,\tau) = (c\tau+d)^{-2}$ in case γ has lower row (c,d).

Recall from §3.4 that the character of $G^{(\ell)}$ attached to its signed permutation representation is denoted $g \mapsto \chi^{(\ell)}_g$ in Tables 14-19 and that of the (unsigned) permutation representation is denoted $g \mapsto \bar{\chi}^{(\ell)}_g$. Define $\chi^{(\ell)}_{g,r}$ for $0 < r < \ell$ by setting $\chi^{(\ell)}_{g,r} = \chi^{(\ell)}_g$ in case r is even and $\chi^{(\ell)}_{g,r} = \bar{\chi}^{(\ell)}_g$ otherwise. Then the shadow of $H^{(\ell)}_g$ is the function $S^{(\ell)}_g = \left(S^{(\ell)}_{g,r}\right)$ with components related to those of $S^{(\ell)}$ by $S^{(\ell)}_{g,r} = \chi^{(\ell)}_{g,r}S^{(\ell)}_r$. In particular, $H^{(\ell)}_g$ is a vector-valued modular form of weight 1/2 for $\Gamma_0(n_g)$ when $\chi^{(\ell)}_g = \bar{\chi}^{(\ell)}_g = 0$.

To summarise, we claim that our proposed McKay–Thompson series $H^{(\ell)}_g$ are such that if we define $\hat{H}^{(\ell)}_g = \left(\hat{H}^{(\ell)}_{g,r}\right)$ by setting

$$\hat{H}^{(\ell)}_{g,r}(\tau) = H_{g,r}(\tau) - \frac{\chi^{(\ell)}_{g,r}}{\sqrt{2\ell}}\frac{1}{(4i)^{1/2}}\int_{-\bar{\tau}}^{i\infty}(z+\tau)^{-1/2}S^{(\ell)}_r(z)\mathrm{d}z \tag{4.46}$$

for $0 < r < \ell$ then $\hat{H}^{(\ell)}_g$ is invariant for the weight 1/2 action of $\Gamma_0(n)$ on $(\ell-1)$-vector-valued functions on $\mathbb{H}$ given by

$$\left(\hat{H}^{(\ell)}_g\Big|_{1/2,n|h}\gamma\right)(\tau) = \nu_{n|h}(\gamma)\hat{H}^{(\ell)}_g(\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{1/4} \tag{4.47}$$

where $n = n_g$ and $h = h_g$ and $\nu_{n|h}$ is defined by (4.44). This statement completely describes the (conjectured) automorphy of the mock modular forms $H^{(\ell)}_g$ (cf. §5.3).

5 Conjectures

We have described the umbral forms $Z^{(\ell)}$, $H^{(\ell)}$ and $\Phi^{(\ell)}$ in §2 and the umbral groups $G^{(\ell)}$ in §3 and we have introduced families $\{H^{(\ell)}_g \mid g \in G^{(\ell)}\}$ of vector-valued mock modular forms in §4. The discussions of those sections clearly demonstrate the distinguished nature of these objects, and we have mentioned some coincidences relating the groups $G^{(\ell)}$ and the forms $H^{(\ell)}_g$ directly. In this section we present, in a more systematic fashion, evidence that the relationship between the $G^{(\ell)}$ and the $H^{(\ell)}$ is more than coincidental. Our observations lead naturally to conjectures that we hope will serve as first steps in revealing the structural nature of the mechanism underlying umbral moonshine.

5.1 Modules

As was mentioned in §4, after comparison of the character tables (cf. §B.1) of the umbral groups $G^{(\ell)}$ with the Fourier coefficient tables (cf. §C) for the forms $H_g^{(\ell)}$ it becomes apparent that the low degree Fourier coefficients of $H^{(\ell)} = H_e^{(\ell)}$ may be interpreted as degrees of representations of $G^{(\ell)}$ in such a way that the corresponding coefficients of $H_g^{(\ell)}$ are recovered by substituting character values at g; we have tabulated evidence for this in the form of explicit combinations of irreducible representations in §D. This observation suggests the existence of bi-graded $G^{(\ell)}$-modules $K^{(\ell)} = \bigoplus K_{r,d}^{(\ell)}$ whose bi-graded dimensions are recovered via Fourier coefficients from the vector-valued mock modular forms $H^{(\ell)} = \left(H_r^{(\ell)}\right)$.

Conjecture 5.1. *We conjecture that for $\ell \in \Lambda = \{2,3,4,5,7,13\}$ there exist naturally defined $\mathbb{Z}\times\mathbb{Q}$-graded $G^{(\ell)}$-modules*

$$K^{(\ell)} = \bigoplus_{\substack{r\in\mathbb{Z}\\0<r<\ell}} K_r^{(\ell)} = \bigoplus_{\substack{r,k\in\mathbb{Z}\\0<r<\ell}} K_{r,k-r^2/4\ell}^{(\ell)} \tag{5.1}$$

such that the graded dimension of $K^{(\ell)}$ is related to the vector-valued mock modular form $H^{(\ell)} = \left(H_r^{(\ell)}\right)$ by

$$H_r^{(\ell)}(\tau) = -2\delta_{r,1}q^{-1/4\ell} + \sum_{\substack{k\in\mathbb{Z}\\r^2-4k\ell<0}} \dim\left(K_{r,k-r^2/4\ell}^{(\ell)}\right) q^{k-r^2/4\ell} \tag{5.2}$$

for $q = \mathrm{e}(\tau)$ and such that the vector-valued mock modular forms $H_g^{(\ell)} = \left(H_{g,r}^{(\ell)}\right)$ described in §4 (and partially in the tables of §C) are recovered from $K^{(\ell)}$ via graded trace functions according to

$$H_{g,r}^{(\ell)}(\tau) = -2\delta_{r,1}q^{-1/4\ell} + \sum_{\substack{k\in\mathbb{Z}\\r^2-4k\ell<0}} \mathrm{tr}_{K_{r,k-r^2/4\ell}^{(\ell)}}(g) q^{k-r^2/4\ell}. \tag{5.3}$$

Remark 5.2. Recall that a *superspace* is a $\mathbb{Z}/2\mathbb{Z}$-graded vector space and if $V = V_{\bar 0}\oplus V_{\bar 1}$ is such an object and $T : V \to V$ is a linear operator preserving the grading then the *supertrace* of T is given by $\mathrm{str}_V T = \mathrm{tr}_{V_{\bar 0}} T - \mathrm{tr}_{V_{\bar 1}} T$ where $\mathrm{tr}_W T$ denotes the usual trace of T on W. Since the coefficient of $q^{-1/4\ell}$ in $H_{g,1}^{(\ell)}$ is $c_{g,1}^{(\ell)}(-1/4\ell) = -2$ for all $g \in G^{(\ell)}$ for all ℓ it is natural to expect that this term may be interpreted as the supertrace of $g \in G^{(\ell)}$ on a trivial $G^{(\ell)}$-supermodule $K_{1,-1/4\ell}^{(\ell)}$ with vanishing even part and 2-dimensional odd part. Thus Conjecture 5.1 implies the existence of a $G^{(\ell)}$-supermodule $K^{(\ell)} = \bigoplus K_{r,d}^{(\ell)}$ such that $K_{r,d}^{(\ell)}$ is purely even or purely odd according as d is positive or negative, and such that $H^{(\ell)}_{g,r} = \sum_k \mathrm{str}_{K^{(\ell)}_{r,k-r^2/4\ell}}(g)q^{k-r^2/r\ell}$ for each $0 < r < \ell$ and $g \in G^{(\ell)}$.

Remark 5.3. For $\ell \in \Lambda$ and $0 < r, s < \ell$ we have that $r^2 \equiv s^2 \pmod{4\ell}$ implies $r = s$ so the first index r in $K^{(\ell)}_{r,d}$ can be deduced from d since it is the unique $0 < r < \ell$ such that $d + r^2/4\ell$ is an integer. Thus we may dispense with the first of the two gradings on the conjectural $G^{(\ell)}$-modules $K^{(\ell)}$ and regard them as $\mathbb{Q}$-graded $K^{(\ell)} = \bigoplus_d K^{(\ell)}_d$ by the rationals of the form $d = n - r^2/4\ell$ for $n \in \mathbb{Z}$ and $0 < r < \ell$ without introducing any ambiguity.

5.2 Moonshine

In the case of monstrous moonshine the McKay–Thompson series T_g for g in the monster group have the astonishing property that they all serve as generators for the function fields of their invariance groups (cf. [1]). In other words, if Γ_g is the subgroup of $PSL_2(\mathbb{R})$ consisting of the isometries $\gamma : \mathbb{H} \to \mathbb{H}$ such that $T_g(\gamma\tau) = T_g(\tau)$ for all $\tau \in \mathbb{H}$ then T_g induces an isomorphism from the compactification of $\Gamma_g\backslash\mathbb{H}$ to the Riemann sphere (being the one point compactification of $\mathbb{C}$). In particular, Γ_g is a genus zero subgroup of $PSL_2(\mathbb{R})$, and so this property is commonly referred to as the *genus zero property* of monstrous moonshine.

By now there are many methods extant for constructing graded vector spaces with algebraic structure whose graded dimensions are modular functions—suitable classes of vertex algebras, for example, serve this purpose (cf. [114, 115])—and so one can expect to obtain analogues of the McKay–Thompson series T_g by equipping such an algebraic structure with the action of a group. But there is no guarantee that such a procedure will result in functions that have the genus zero property of monstrous moonshine, and so it is this genus zero property that distinguishes the observations of McKay, Thompson, Conway and Norton regarding the Monster group from any number of more generic connections between finite groups and modular functions. In what follows we will propose a conjecture that may be regarded as the natural analogue of the Conway–Norton conjecture for umbral moonshine.

Suppose that $T : \mathbb{H} \to \mathbb{C}$ is a holomorphic function with invariance group $\Gamma < PSL_2(\mathbb{R})$ and a simple pole in $q = e(\tau)$ as $\tau \to i\infty$. Suppose also that Γ is commensurable with $PSL_2(\mathbb{Z})$ and that the *translation subgroup* Γ_∞, consisting of the elements of Γ with upper-triangular preimages in $SL_2(\mathbb{R})$, is generated by $\tau \mapsto \tau + 1$. It is shown in [25] that such a function T is a generator for the field of Γ-invariant functions on $\mathbb{H}$ if and only if $T(\tau)$ coincides with the weight

0 *Rademacher sum*

$$R_\Gamma(\tau) = \operatorname{Reg}\left(\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} q^{-1}\big|_0\,\gamma\right) \tag{5.4}$$

attached to Γ. The sum here is over representatives for the cosets of Γ_∞ in Γ and $f \mapsto f|_0\gamma$ denotes the weight 0 action of γ on holomorphic functions (viz., $(f|_0\gamma)(\tau) = f(\gamma\tau)$). We write $\operatorname{Reg}(\cdot)$ to indicate a regularisation procedure first realised by Rademacher (for the case that $\Gamma = PSL_2(\mathbb{Z})$) in [26]. According then to the result of [25] the genus zero property of monstrous moonshine may be reformulated in the following way:

> For each element g in the monster group the McKay–Thompson series T_g satisfies $T_g = R_{\Gamma_g}$ when Γ_g is the invariance group of T_g.

This may be compared with the article [24] which considers the mock modular forms $H_g^{(2)}$ attached to the largest Mathieu group $G^{(2)} \simeq M_{24}$ via the observation of Eguchi–Ooguri–Tachikawa and applies some of the philosophy of [25] to the problem of finding a uniform construction for these functions. The solution developed in [24] is that the function $H_g^{(2)}$ coincides with the weight 1/2 Rademacher sum

$$R^{(2)}_{n|h}(\tau) = \operatorname{Reg}\left(\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(n)} -2q^{-1/8}\Big|_{1/2,n|h}\,\gamma\right) \tag{5.5}$$

where $n = n_g$ and $h = h_g$ are integers determined by the defining permutation representation of M_{24} (cf. Table 14) and $f \mapsto f|_{1/2,n|h}\gamma$ denotes a certain weight 1/2 action of $\Gamma_0(n)$ determined by n and h. (We refer to [24] for full details.) By comparison with the previous paragraph we thus arrive at a direct analogue of the genus zero property of monstrous moonshine that holds for all the umbral forms with $\ell = 2$:

> For each element g in the largest Mathieu group the McKay–Thompson series $H_g^{(2)}$ satisfies $H_g^{(2)} = R^{(2)}_{n|h}$ when $n = n_g$ and $h = h_g$.

We conjecture that this genus zero property extends to all the functions of umbral moonshine.

Conjecture 5.4 (umbral moonshine)**.** *We conjecture that for each $\ell \in \Lambda$ and each $g \in G^{(\ell)}$ we* *have* $H_g^{(\ell)} = R_{n|h}^{(\ell)}$ *where* $n = n_g$ *and* $h = h_g$ *are as specified in* §B.2 *and*

$$R_{n|h}^{(\ell)} = \operatorname{Reg}\left(\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(n)} \begin{pmatrix} -2q^{-1/4\ell} \\ 0 \\ \vdots \\ 0 \end{pmatrix}\Bigg|_{1/2,n|h} \gamma \right) \tag{5.6}$$

is a vector-valued generalisation of the Rademacher sum $R_{n|h}^{(2)}$ *adapted to the weight* $1/2$ *action of* $\Gamma_0(n)$ *on* $(\ell-1)$*-vector-valued functions on* $\mathbb{H}$ *that is defined in* (4.47).

5.3 Modularity

A beautiful feature of the umbral moonshine conjecture is that it implies the precise nature of the modularity of the Thompson series $H_g^{(\ell)}$. We record these implications explicitly as conjectures in this short section.

Conjecture 5.5. *We conjecture that the graded supertrace functions* (5.3) *for fixed* $\ell \in \Lambda$ *and* $g \in G^{(\ell)}$ *and varying* $0 < r < \ell$ *define the components of a vector-valued mock modular form* $H_g^{(\ell)}$ *of weight* $1/2$ *on* $\Gamma_0(n_g)$ *with shadow function* $S_g^{(\ell)} = (S_{g,r}^{(\ell)}) = (\chi_{g,r}^{(\ell)} S_r^{(\ell)})$ *where* $S^{(\ell)} = (S_r^{(\ell)})$ *is the vector-valued theta series described in* §2.3*, the* $\chi_{g,r}^{(\ell)}$ *are determined from the twisted Euler characters of* $G^{(\ell)}$ *(cf.* §B.2*) by* $\chi_{g,r}^{(\ell)} = \bar{\chi}_g^{(\ell)}$ *for* r *odd and* $\chi_{g,r}^{(\ell)} = \chi_g^{(\ell)}$ *for* r *even, and* n_g *denotes the order of the image of* $g \in G^{(\ell)}$ *in the factor group* $\bar{G}^{(\ell)}$ *(cf.* §3.1,3.4*).*

Recall from §4.8 that for $\ell \in \Lambda$ and a pair (n, h) of positive integers we have the matrix-valued function $\nu_{n|h}^{(\ell)}$ on $\Gamma_0(n)$ defined in (4.44). Recall also that we have attached a pair (n_g, h_g) to each $g \in G^{(\ell)}$ for each $\ell \in \Lambda$ in §4.8.

Conjecture 5.6. *We conjecture that the multiplier system of* $H_g^{(\ell)}$ *is given by* $\nu_{n|h}^{(\ell)}$ *when* $n = n_g$ *and* $h = h_g$ *for all* $g \in G^{(\ell)}$ *for all* $\ell \in \Lambda$*.*

5.4 Discriminants

One of the most striking features of umbral moonshine is the apparently intimate relation between the number fields on which the irreducible representations of $G^{(\ell)}$ are defined and the discriminants of the vector-valued mock modular form $H^{(\ell)}$. We discuss the evidence for this relation and formulate conjectures about it in this section.

First we observe that the discriminants of the components $H_r^{(\ell)}$ of the mock modular form $H^{(\ell)} = H_e^{(\ell)}$ determine some important properties of the representations of $G^{(\ell)}$, where we say that an integer D is a *discriminant of* $H^{(\ell)}$ if there exists a term $q^d = q^{-\frac{D}{4\ell}}$ with non-vanishing Fourier coefficient in at least one of the components. The following result can be verified explicitly using the tables in §§B,C.

Proposition 5.7. *Let $\ell \in \Lambda$. If $n > 1$ is an integer satisfying*

1. *there exists an element of $G^{(\ell)}$ of order n, and*
2. *there exists an integer λ that is co-prime to n such that $D = -n\lambda^2$ is a discriminant of $H^{(\ell)}$,*

then there exists at least one pair of irreducible representations ϱ and ϱ^* of $G^{(\ell)}$ and at least one element $g \in G^{(\ell)}$ such that $\mathrm{tr}_\varrho(g)$ is not rational but

$$\mathrm{tr}_\varrho(g), \mathrm{tr}_{\varrho^*}(g) \in \mathbb{Q}(\sqrt{-n}) \tag{5.7}$$

and n divides $o(g)$.

The finite list of integers n satisfying the two conditions of Proposition 5.7 is given in Table 7.

From now on we say that an irreducible representation ϱ of the umbral group $G^{(\ell)}$ is of *type* n if n is an integer satisfying the two conditions of Proposition 5.7 and the character values of ϱ generate the field $\mathbb{Q}(\sqrt{-n})$. Evidently, irreducible representations of type n come in pairs (ϱ, ϱ^*) with $\mathrm{tr}_{\varrho^*}(g)$ the complex conjugate of $\mathrm{tr}_\varrho(g)$ for all $g \in G^{(\ell)}$. The list of all irreducible representations of type n is given in Table 7. (See §B.1 for the character tables of the $G^{(\ell)}$ and our notation for irreducible representations.)

Recall that the *Frobenius–Schur indicator* of an irreducible ordinary representation of a finite group is 1, -1 or 0 according as the representation admits an invariant symmetric bilinear form, an invariant skew-symmetric bilinear form, or no invariant bilinear form, respectively. The representations admitting no invariant bilinear form are precisely those whose character values are not all real. We can now state the next observation.

Proposition 5.8. *For each $\ell \in \Lambda$ an irreducible representation ϱ of $G^{(\ell)}$ has Frobenius–Schur indicator* 0 *if and only if it is of type n for some n.*

The *Schur index* of an irreducible representation ϱ of a finite group G is the smallest positive integer s such that there exists a degree s extension k of the field generated by the character values $\mathrm{tr}_\varrho(g)$ for $g \in G$ such that ϱ can be realised over k. Inspired by Proposition 5.8 we make the following conjecture.

Conjecture 5.9. *If ϱ is an irreducible representation of $G^{(\ell)}$ of type n then the Schur index of ϱ is equal to* 1*.*

In other words, we conjecture that the irreducible $G^{(\ell)}$-representations of type n can be realised over $\mathbb{Q}(\sqrt{-n})$. For $\ell = 2$ this speculation is in fact a theorem, since it is known [116] that the Schur indices for M_{24} are always 1. For $\ell = 3$ it is also known [116] that the Schur indices for $\bar{G}^{(3)} = M_{12}$ are also always 1. Moreover, the representations of $G^{(3)} \simeq 2.\bar{G}^{(3)}$ with characters χ_{16} and χ_{17} in the notation of Table 9 have been constructed explicitly over $\mathbb{Q}(\sqrt{-2})$ in [117]. Finally, Proposition 5.8 constitutes a non-trivial consistency check for Conjecture 5.9 since the Schur index is at least 2 for a representation with Frobenius–Schur indicator equal to -1.

ℓ	n	(ϱ, ϱ^*)
2	$7, 15, 23$	$(\chi_3, \chi_4), (\chi_5, \chi_6), (\chi_{10}, \chi_{11}), (\chi_{12}, \chi_{13}), (\chi_{15}, \chi_{16})$
3	$5, 8, 11, 20$	$(\chi_4, \chi_5), (\chi_{16}, \chi_{17}), (\chi_{20}, \chi_{21}), (\chi_{22}, \chi_{23}), (\chi_{25}, \chi_{26})$
4	$3, 7$	$(\chi_2, \chi_3), (\chi_{13}, \chi_{14}), (\chi_{15}, \chi_{16})$
5	4	$(\chi_8, \chi_9), (\chi_{10}, \chi_{11}), (\chi_{12}, \chi_{13})$
7	3	$(\chi_2, \chi_3), (\chi_6, \chi_7)$
13	4	(χ_3, χ_4)

Table 7: The irreducible representations of type n.

Armed with the preceding discussion we are now ready to state our main observation for the discriminant property of umbral moonshine. For the purpose of stating this we temporarily write $K^{(\ell)}_{r,d}$ for the ordinary representation of $G^{(\ell)}$ with character $g \mapsto c^{(\ell)}_{g,r}(d)$ where the coefficients $c^{(\ell)}_{g,r}(d)$ are assumed to be those given in §C.

Proposition 5.10. *Let n be one of the integers in Table 7 and let λ_n be the smallest positive integer such that $D = -n\lambda_n^2$ is a discriminant of $H^{(\ell)}$. Then $K^{(\ell)}_{r,-D/4\ell} = \varrho_n \oplus \varrho_n^*$ where ϱ_n and ϱ_n^* are dual irreducible representations of type n. Conversely, if ϱ is an irreducible representation of type n and $-D$ is the smallest positive integer such that $K^{(\ell)}_{r,-D/4\ell}$ has ϱ as an irreducible constituent then there exists an integer λ such that $D = -n\lambda^2$.*

Extending this we make the following conjecture.

Conjecture 5.11. *If D is a discriminant of $H^{(\ell)}$ which satisfies $D = -n\lambda^2$ for some integer λ then the representation $K^{(\ell)}_{r,-D/4\ell}$ has at least one dual pair of irreducible representations of type n arising as irreducible constituents.*

We conclude this section with conjectures arising from the observation (cf. §D) that the conjectural $G^{(\ell)}$-module $K^{(\ell)}_{r,d}$ is typically isomorphic to several copies of a single representation. Say a G-module V is a *doublet* if it is isomorphic to the direct sum of two copies of a single representation of G.

Conjecture 5.12. *For $\ell \in \Lambda = \{2,3,4,5,7,13\}$ the representation $K^{(\ell)}_{r,-D/4\ell}$ is a doublet if and only if $D \neq -n\lambda^2$ for any integer λ for any n satisfying the conditions of Proposition 5.7.*

To see some evidence for Conjecture 5.12 one can inspect the proposed decompositions of the representations $K^{(\ell)}_{r,d}$ in the tables in §D for the following discriminants:

- $-D = 7, 15, 23, 63, 135, 175, 207$ for $\ell = 2$,
- $-D = 8, 11, 20, 32, 44, 80$ for $\ell = 3$,
- $-D = 7, 12, 28, 63, 108$ for $\ell = 4$,
- $-D = 4, 16, 64, 144, 196$ for $\ell = 5$,
- $-D = 3\lambda^2$, $\lambda = 1, \ldots, 9$, $\lambda \neq 7$ for $\ell = 7$,
- $-D = 4\lambda^2$, $\lambda = 1, \ldots, 11$ for $\ell = 13$.

5.5 Geometry and Physics

Beyond the conjectures already mentioned above several interesting and important questions remain regarding the structural nature of umbral moonshine. In the case that $\ell = 2$ there are strong indications that a deep relationship to $K3$ surfaces—extending in some way the relation [118, 119] between finite groups of $K3$ surface symplectomorphisms and subgroups of M_{23}—is responsible for the relationship between $G^{(2)} \simeq M_{24}$ and $H^{(2)}$. One such indication is the fact that the Jacobi form $Z^{(2)}$, from which $H^{(2)}$ may be obtained by decomposing with respect to superconformal characters, coincides with the elliptic genus of a(ny) complex $K3$ surface. It is natural then to ask if there are analogous geometric interpretations for the remaining extremal Jacobi forms $Z^{(\ell)}$ for $\ell \in \Lambda$, and a positive answer to this question will be a first step in determining the geometric significance of the umbral groups $G^{(\ell)}$ and the attached mock modular forms $H^{(\ell)}_g$.

In a series of papers [120, 121, 88, 87, 89] Gritsenko–Nikulin develop applications of Siegel modular forms to mirror symmetry for $K3$ surfaces (cf. [122]). Many of the Siegel forms arising are realised as additive or exponential lifts of (weak) Jacobi forms, and amongst many examples the exponential lifts $\Phi^{(\ell)}$ of the particular forms $Z^{(\ell)}$ for $\ell \in \{2,3,4,5\} \subset \Lambda$ appear in connection with explicitly defined families of lattice polarised $K3$ surfaces in [87]. It is natural to ask if there is an analogous relationship between the umbral $Z^{(\ell)}$ and suitably defined mirror families of $K3$ surfaces for the remaining ℓ in Λ.

As mentioned in the introduction, monstrous moonshine involves aspects of conformal field theory and string theory and there are several hints that umbral moonshine will also play a role in string theory. The most obvious hint at $\ell = 2$ is through the fact mentioned above that the weak Jacobi form $Z^{(2)}$ coincides with the elliptic genus of a $K3$ surface and the fact that $K3$ surfaces play a prominent role in the study of superstring compactification. The Siegel form $\Phi^{(2)} = (\Delta_5)^2$ which is the multiplicative lift of the weak Jacobi form $Z^{(2)}$ also plays a distinguished role in type II string theory on $K3 \times E$ where E is an elliptic curve and $1/\Phi^{(2)}$ occurs as the generating function which counts the number of 1/4 BPS dyon states [123]. BPS states in a supersymmetric theory are states that are annihilated by some of the supercharges of the theory; the 1/4 BPS states of interest here are annihilated by 4 of the 16 supercharges of a theory with $N = 4$ supersymmetry in four spacetime dimensions and are termed *dyons* because they necessarily carry both electric and magnetic charges. The relation between the zeros of $\Phi^{(2)}$ and the wall-crossing phenomenon in this theory has been studied in [124, 125, 126, 127] and this relation leads to a connection between the counting of BPS black hole states in string theory and mock modular forms [45]. For a pedagogical review of this material see [128, 38].

There are several hints that umbral moonshine for $\ell > 2$ will also play a role in string theory. The first of these occurs in the study of $N = 2$ dual pairs of string theories [129, 130]. An $N = 2$ dual pair consists of a compactification of the heterotic string on a product of a $K3$ surface with an elliptic curve E with a specific choice of $E_8 \times E_8$ gauge bundle on $K3 \times E$ and a conjectured dual description of this model in terms of IIA string theory on a Calabi–Yau threefold X which admits a $K3$ fibration. The low-energy description of such theories involves a vector-multiplet moduli space which is a special Kähler manifold of the form

$$\mathcal{M}_{vm}^{s+2,2} = \frac{SU(1,1)}{U(1)} \times \mathcal{N}^{s+2,2} \tag{5.8}$$

where $\mathcal{N}^{s+2,2}$ is the quotient $\Gamma\backslash\mathcal{H}^{s+1,1}$ of the generalised upper half-plane

$$\mathcal{H}^{s+1,1} = O(s+2,2;\mathbb{R})/(O(s+2) \times O(2)) \tag{5.9}$$

by an arithmetic subgroup Γ in $SO(s+2,2;\mathbb{R})$, with Γ depending on the specific model in question. One-loop string computations in the heterotic string lead to automorphic forms on $\mathcal{N}^{s+2,2}$ via a generalised theta lift constructed in [131, 132], see [133] for a review. In models with $s = 1$ this leads to automorphic forms on $\Gamma\backslash O(3,2;\mathbb{R})/(O(3) \times O(2))$, that is to Siegel modular forms on the genus 2 Siegel upper half-space $\mathbb{H}_2 \cong \mathcal{H}^{2,1}$.

A particular example of such an $N = 2$ dual pair was studied in [134, 135] with $s = 1$ and involving a dual description in terms of Type II string theory on a Calabi–Yau 3-fold with Hodge numbers $(h^{1,1}, h^{2,1}) = (4, 148)$. In the heterotic description the arithmetic subgroup which appears is the paramodular group Γ_2, the Siegel modular form which occurs is the exponential lift of $Z^{(3)}$ described in §2.6 and the Calabi–Yau 3-fold is also elliptically fibered and the elliptic fiber is of E_7 type. We thus see many of the ingredients of $\ell = 3$ umbral moonshine appearing in the context of a specific dual pair of string theories with $N = 2$ spacetime supersymmetry. It will be very interesting to investigate whether other aspects of $\ell = 3$ umbral moonshine can be realized in these models and whether $N = 2$ dual pairs exist which exhibit elements of umbral moonshine for $\ell > 3$.

A second way that elements of umbral moonshine for $\ell > 2$ are likely to appear in string theory involves a generalisation of the generating function discussed above that counts 1/4 BPS states. Type II string theory on $K3 \times E$ preserves $N = 4$ spacetime supersymmetry. There exist orbifold versions of this model that also preserve $N = 4$ spacetime supersymmetry that are known in the physics literature as CHL models. To construct such a model one chooses a $K3$ surface with a $\mathbb{Z}/m\mathbb{Z}$ hyper-Kähler automorphism and constructs the orbifold theory $(K3 \times E)/(\mathbb{Z}/m\mathbb{Z})$ where $\mathbb{Z}/m\mathbb{Z}$ is a freely acting symmetry realised as the product of a hyper-Kähler automorphism of the $K3$ surface and an order m translation along E. For $m = 2, 3, 4$ it is possible to find $K3$ surfaces with $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$ symmetry, and in this case one can construct a CHL model that utilises the first $\mathbb{Z}/m\mathbb{Z}$ factor in the orbifold construction and has the second $\mathbb{Z}/m\mathbb{Z}$ factor acting as a symmetry which preserves the holomorphic 3-form of the Calabi–Yau space $(K3 \times E)/(\mathbb{Z}/m\mathbb{Z})$. It was proposed in [136] that the generating function which counts 1/4 BPS states weighted by an element of $\mathbb{Z}/m\mathbb{Z}$ is the reciprocal of the Siegel form $\Phi^{(m+1)}$ for $m \in \{2, 3, 4\}$. In both this construction and in the study of $N = 2$ dual pairs the appearance of some of the umbral Siegel forms $\Phi^{(\ell)}$ can be anticipated, although many details remain to be worked out. The action of the umbral groups $G^{(\ell)}$ remains more elusive, and deeper insight into possible connections between string theory and umbral moonshine will undoubtedly require progress in understanding the actions of these groups in terms of their action on BPS states.

We plan to elaborate further on the topics mentioned above in future work.

Acknowledgements

We thank Chris Cummins, Atish Dabholkar, Noam Elkies, Igor Frenkel, George Glauberman, Ian Grojnowski, Kobi Kremnizer, Anthony Licata, Greg Moore, Sameer Murthy, Ken Ono, Cumrun Vafa, Erik Verlinde, Edward Witten, Shing-Tung Yau and Don Zagier for helpful comments and discussions. We thank Noam Elkies and Ken Ono in particular for communication regarding the non-vanishing of critical central values for modular L-functions.

We are especially grateful to Don Zagier for showing us his computation of the first few Fourier coefficients of the functions (that we have here denoted) $H^{(\ell)}$ for $\ell \in \{2,3,4,5,7\}$ at the "Workshop on Mathieu Moonshine" held in July 2011 at the Swiss Federal Institute of Technology (ETH), Zürich. This event was an important catalyst for precipitating the idea that there could be a family of groups attached to a distinguished family of vector-valued modular forms, in a way that generalises the connection between M_{24} and the (scalar-valued) mock modular form $H^{(2)}$. We are also grateful to Kazuhiro Hikami for his talk at the ETH meeting on character decompositions of elliptic genera, and for kindly sharing his slides from that presentation. The analysis described therein proved useful for us in developing a computational approach to approximate Fourier coefficients of the $H^{(\ell)}$ and this furnished an important source of experimental evidence to guide our umbral explorations. We warmly thank the organisers of the ETH meeting for making these exchanges possible and for providing an excellent workshop experience more generally.

The research of MC is partially supported by the NSF grant FRG 0854971. The work of JH was supported by the NSF grant 0855039, the hospitality of the Aspen Center for Physics, the LPTHE of the Université Pierre et Marie Curie, and the Isaac Newton Institute for Mathematical Sciences. We also thank George Armhold for instruction on the efficient implementation of Pari in OS X, and Jia-Chen Fu for the provision of computational resources at a crucial stage in our investigations.

We are grateful to George Glauberman, Sander Mack-Crane, and the referee, for pointing out errors in an earlier draft. We are grateful to the referee also for suggesting many valuable improvements to the text.

A Modular Forms

A.1 Dedekind Eta Function

The *Dedekind eta function*, denoted $\eta(\tau)$, is a holomorphic function on the upper half-plane defined by the infinite product

$$\eta(\tau)=q^{1/24}\prod_{n\geq 1}(1-q^n)$$

where $q=\mathrm{e}(\tau)=e^{2\pi\mathbf{i}\tau}$. It is a modular form of weight 1/2 for the modular group $SL_2(\mathbb{Z})$ with multiplier $\epsilon: SL_2(\mathbb{Z})\to\mathbb{C}^*$, which means that

$$\epsilon(\gamma)\eta(\gamma\tau)\mathrm{jac}(\gamma,\tau)^{1/4}=\eta(\tau)$$

for all $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in SL_2(\mathbb{Z})$, where $\mathrm{jac}(\gamma,\tau)=(c\tau+d)^{-2}$. The *multiplier system* ϵ may be described explicitly as

$$\epsilon\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{cases}\mathrm{e}(-b/24), & c=0,\, d=1\\ \mathrm{e}(-(a+d)/24c+s(d,c)/2+1/8), & c>0\end{cases}\tag{A.1}$$

where $s(d,c)=\sum_{m=1}^{c-1}(d/c)((md/c))$ and $((x))$ is 0 for $x\in\mathbb{Z}$ and $x-\lfloor x\rfloor-1/2$ otherwise. We can deduce the values $\epsilon(a,b,c,d)$ for $c<0$, or for $c=0$ and $d=-1$, by observing that $\epsilon(-\gamma)=\epsilon(\gamma)\,\mathrm{e}(1/4)$ for $\gamma\in SL_2(\mathbb{Z})$.

Let T denote the element of $SL_2(\mathbb{Z})$ such that $\mathrm{tr}(T)=2$ and $T\tau=\tau+1$ for $\tau\in\mathbb{H}$. Observe that

$$\epsilon(T^m\gamma)=\epsilon(\gamma T^m)=\mathrm{e}(-m/24)\epsilon(\gamma)$$

for $m\in\mathbb{Z}$.

A.2 Jacobi Theta Functions

We define the *Jacboi theta functions* $\theta_i(\tau,z)$ as follows for $q=e(\tau)$ and $y=e(z)$.

$$\theta_1(\tau,z)=-iq^{1/8}y^{1/2}\prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^{n-1})$$

$$\theta_2(\tau,z)=q^{1/8}y^{1/2}\prod_{n=1}^{\infty}(1-q^n)(1+yq^n)(1+y^{-1}q^{n-1})$$

$$\theta_3(\tau,z) = \prod_{n=1}^{\infty}(1-q^n)(1+y\,q^{n-1/2})(1+y^{-1}q^{n-1/2})$$

$$\theta_4(\tau,z) = \prod_{n=1}^{\infty}(1-q^n)(1-y\,q^{n-1/2})(1-y^{-1}q^{n-1/2})$$

Note that there are competing conventions for $\theta_1(\tau,z)$ in the literature and our normalisation may differ from another by a factor of -1 (or possibly $\pm i$).

A.3 Higher Level Modular Forms

The congruence subgroups of the modular group $SL_2(\mathbb{Z})$ that are most relevant for this paper are

$$\Gamma_0(N) \quad = \quad \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}), c = 0 \bmod N \right\}. \tag{A.2}$$

For $N>1$ a (non-zero) modular form of weight 2 for $\Gamma_0(N)$ is given by

$$\begin{aligned} \Lambda_N(\tau) &= N\,q\partial_q \log\left(\frac{\eta(N\tau)}{\eta(\tau)}\right) \\ &= \frac{N(N-1)}{24}\left(1+\frac{24}{N-1}\sum_{k>0}\sigma(k)(q^k - Nq^{Nk})\right) \end{aligned} \tag{A.3}$$

where $\sigma(k)$ is the divisor function $\sigma(k)=\sum_{d|k} d$.

Observe that a modular form on $\Gamma_0(N)$ is a modular form on $\Gamma_0(M)$ whenever $N|M$, and for some small N the space of forms of weight 2 is spanned by the $\Lambda_d(\tau)$ for d a divisor of N. By contrast, in the case that $N=11$ we have the *newform*

$$f_{11}(\tau) = \eta^2(\tau)\eta^2(11\tau) \tag{A.4}$$

which is a cusp form of weight 2 for $\Gamma_0(11)$ that is not a multiple of $\Lambda_{11}(\tau)$. We meet the newforms

$$f_{14}(\tau) \quad = \quad \eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau,) \tag{A.5}$$

$$f_{15}(\tau) \quad = \quad \eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau), \tag{A.6}$$

$$f_{20}(\tau) \quad = \quad \eta(2\tau)^2\eta(10\tau)^2, \tag{A.7}$$

at $N=14$, $N=15$ and $N=20$, respectively, and together with f_N the functions $\Lambda_d(\tau)$ for $d|N$

span the space of weight 2 forms on $\Gamma_0(N)$ for $N = 11, 14, 15$.

For $N = 23$ there is a two dimensional space of newforms. We may use the basis

$$\begin{aligned} f_{23,a}(\tau) &= \frac{\eta(\tau)^3\eta(23\tau)^3}{\eta(2\tau)\eta(46\tau)} + 3\eta(\tau)^2\eta(23\tau)^2 + 4\eta(\tau)\eta(2\tau)\eta(23\tau)\eta(46\tau) + 4\eta(2\tau)^2\eta(46\tau)^2 \\ f_{23,b}(\tau) &= \eta(\tau)^2\eta(23\tau)^2 \end{aligned} \tag{A.8}$$

satisfying $f_{23,a} = q + O(q^3)$ and $f_{23,b} = q^2 + O(q^3)$. (The normalised Hecke-eigenforms of weight 2 for $\Gamma_0(23)$ are $f_{23,a} - \frac{1}{2}(1 \pm \sqrt{5})f_{23,b}$.)

For $N = 44$ there is a unique newform up to a multiplicative constant. It satisfies

$$f_{44}(\tau) = q + q^3 - 3q^5 + 2q^7 - 2q^9 - q^{11} - 4q^{13} - 3q^{15} + 6q^{17} + 8q^{19} + 2q^{21} - 3q^{23} + 4q^{25} - 5q^{27} + O(q^{28}).$$

See [137] and Chapter 4.D of [138] for more details. A discussion of the ring of weak Jacobi forms of higher level can be found in [139].

B Characters

In §B.1 we give character tables (with power maps and Frobenius–Schur indicators) for each group $G^{(\ell)}$. These were computed with the aid of the computer algebra package GAP4 [98] using the explicit presentations for the $G^{(\ell)}$ that appear in §3.4. We use the abbreviations $a_n = \sqrt{-n}$ and $b_n = (-1 + \sqrt{-n})/2$ in these tables.

The tables in §B.2 furnish the Frame shapes $\Pi_g^{(\ell)}$ and character values $\chi_g^{(\ell)}$ attached to the signed permutation representations (cf. §3.2) of the groups $G^{(\ell)}$ given in §3.4. These Frame shapes and character values can easily be computed by hand; we detail them here explicitly since they can be used to define symbols $n_g|h_g$ which encode the automorphy of the vector-valued mock modular forms $H_g^{(\ell)}$ according to the prescription of §4.8. The symbols $n_g|h_g$ are given in the rows labelled Γ_g in §B.2.

B.1 Irreducible Characters

Table 8: Character table of $G^{(2)} \simeq M_{24}$

$[g]$	FS	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	23A	23B
$[g^2]$		1A	1A	1A	3A	3B	2A	2A	2B	5A	3A	3B	7A	7B	4B	5A	11A	6A	6B	7A	7B	15A	15B	21A	21B	23A	23B
$[g^3]$		1A	2A	2B	1A	1A	4A	4B	4C	5A	2A	2B	7B	7A	8A	10A	11A	4A	4C	14B	14A	5A	5A	7B	7A	23A	23B
$[g^5]$		1A	2A	2B	3A	3B	4A	4B	4C	1A	6A	6B	7B	7A	8A	2B	11A	12A	12B	14B	14A	3A	3A	21B	21A	23B	23A
$[g^7]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	1A	1A	8A	10A	11A	12A	12B	2A	2A	15B	15A	3B	3B	23B	23A
$[g^{11}]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	1A	12A	12B	14A	14B	15B	15A	21A	21B	23B	23A
$[g^{23}]$		1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7A	7B	8A	10A	11A	12A	12B	14A	14B	15A	15B	21A	21B	1A	1A
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	+	23	7	-1	5	-1	-1	3	-1	3	1	-1	2	2	1	-1	1	-1	-1	0	0	0	0	-1	-1	0	0
χ_3	○	45	-3	5	0	3	-3	1	1	0	0	-1	b_7	$\overline{b_7}$	-1	0	1	0	1	$-b_7$	$-\overline{b_7}$	0	0	b_7	$\overline{b_7}$	-1	-1
χ_4	○	45	-3	5	0	3	-3	1	1	0	0	-1	$\overline{b_7}$	b_7	-1	0	1	0	1	$-\overline{b_7}$	$-b_7$	0	0	$\overline{b_7}$	b_7	-1	-1
χ_5	○	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	b_{15}	$\overline{b_{15}}$	0	0	1	1
χ_6	○	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	$\overline{b_{15}}$	b_{15}	0	0	1	1
χ_7	+	252	28	12	9	0	4	4	0	2	1	0	0	0	0	2	-1	1	0	0	0	-1	-1	0	0	-1	-1
χ_8	+	253	13	-11	10	1	-3	1	1	3	-2	1	1	1	-1	-1	0	0	1	-1	-1	0	0	1	1	0	0
χ_9	+	483	35	3	6	0	3	3	3	-2	2	0	0	0	-1	-2	-1	0	0	0	0	1	1	0	0	0	0
χ_{10}	○	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	b_{23}	$\overline{b_{23}}$
χ_{11}	○	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	$\overline{b_{23}}$	b_{23}
χ_{12}	○	990	-18	-10	0	3	6	2	-2	0	0	-1	b_7	$\overline{b_7}$	0	0	0	0	1	b_7	$\overline{b_7}$	0	0	b_7	$\overline{b_7}$	1	1
χ_{13}	○	990	-18	-10	0	3	6	2	-2	0	0	-1	$\overline{b_7}$	b_7	0	0	0	0	1	$\overline{b_7}$	b_7	0	0	$\overline{b_7}$	b_7	1	1
χ_{14}	+	1035	27	35	0	6	3	-1	3	0	0	2	-1	-1	1	0	1	0	0	-1	-1	0	0	-1	-1	0	0
χ_{15}	○	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2b_7$	$2\overline{b_7}$	-1	0	1	0	-1	0	0	0	0	$-b_7$	$-\overline{b_7}$	0	0
χ_{16}	○	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2\overline{b_7}$	$2b_7$	-1	0	1	0	-1	0	0	0	0	$-\overline{b_7}$	$-b_7$	0	0
χ_{17}	+	1265	49	-15	5	8	-7	1	-3	0	1	0	-2	-2	1	0	0	-1	0	0	0	0	0	1	1	0	0
χ_{18}	+	1771	-21	11	16	7	3	-5	-1	1	0	-1	0	0	-1	1	0	0	-1	0	0	1	1	0	0	0	0
χ_{19}	+	2024	8	24	-1	8	8	0	0	-1	-1	0	1	1	0	-1	0	-1	0	1	1	-1	-1	1	1	0	0
χ_{20}	+	2277	21	-19	0	6	-3	1	-3	-3	0	2	2	2	-1	1	0	0	0	0	0	0	0	-1	-1	0	0
χ_{21}	+	3312	48	16	0	-6	0	0	0	-3	0	-2	1	1	0	1	1	0	0	-1	-1	0	0	1	1	0	0
χ_{22}	+	3520	64	0	10	-8	0	0	0	0	-2	0	-1	-1	0	0	0	0	0	1	1	0	0	-1	-1	1	1
χ_{23}	+	5313	49	9	-15	0	1	-3	-3	3	1	0	0	0	-1	-1	0	1	0	0	0	0	0	0	0	0	0
χ_{24}	+	5544	-56	24	9	0	-8	0	0	-1	1	0	0	0	0	-1	0	1	0	0	0	-1	-1	0	0	1	1
χ_{25}	+	5796	-28	36	-9	0	-4	4	0	1	-1	0	0	0	0	1	-1	-1	0	0	0	1	1	0	0	0	0
χ_{26}	+	10395	-21	-45	0	0	3	-1	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	-1	-1

Table 9: Character table of $G^{(3)} \simeq 2.M_{12}$

$[g]$	FS	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8A	8B	8C	8D	20A	20B	11A	22A	11B	22B
$[g^2]$		1A	1A	2A	1A	1A	3A	3A	3B	3B	2B	2B	5A	5A	6B	3A	3A	4B	4B	4C	4C	10A	10A	11B	11B	11A	11A
$[g^3]$		1A	2A	4A	2B	2C	1A	2A	1A	2A	4B	4C	5A	10A	4A	2B	2C	8A	8B	8C	8D	20A	20B	11A	22A	11B	22B
$[g^5]$		1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	1A	2A	12A	6C	6D	8B	8A	8D	8C	4A	4A	11A	22A	11B	22B
$[g^{11}]$		1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8A	8B	8C	8D	20B	20A	1A	2A	1A	2A
χ_1	$+$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	$+$	11	11	-1	3	3	2	2	-1	-1	-1	3	1	1	-1	0	0	-1	-1	1	1	-1	-1	0	0	0	0
χ_3	$+$	11	11	-1	3	3	2	2	-1	-1	3	-1	1	1	-1	0	0	1	1	-1	-1	-1	-1	0	0	0	0
χ_4	$\circ$	16	16	4	0	0	-2	-2	1	1	0	0	1	1	1	0	0	0	0	0	0	-1	-1	b_{11}	b_{11}	$\overline{b_{11}}$	$\overline{b_{11}}$
χ_5	$\circ$	16	16	4	0	0	-2	-2	1	1	0	0	1	1	1	0	0	0	0	0	0	-1	-1	$\overline{b_{11}}$	$\overline{b_{11}}$	b_{11}	b_{11}
χ_6	$+$	45	45	5	-3	-3	0	0	3	3	1	1	0	0	-1	0	0	-1	-1	-1	-1	0	0	1	1	1	1
χ_7	$+$	54	54	6	6	6	0	0	0	0	2	2	-1	-1	0	0	0	0	0	0	0	1	1	-1	-1	-1	-1
χ_8	$+$	55	55	-5	7	7	1	1	1	1	-1	-1	0	0	1	1	1	-1	-1	-1	-1	0	0	0	0	0	0
χ_9	$+$	55	55	-5	-1	-1	1	1	1	1	3	-1	0	0	1	-1	-1	-1	-1	1	1	0	0	0	0	0	0
χ_{10}	$+$	55	55	-5	-1	-1	1	1	1	1	-1	3	0	0	1	-1	-1	1	1	-1	-1	0	0	0	0	0	0
χ_{11}	$+$	66	66	6	2	2	3	3	0	0	-2	-2	1	1	0	-1	-1	0	0	0	0	1	1	0	0	0	0
χ_{12}	$+$	99	99	-1	3	3	0	0	3	3	-1	-1	-1	-1	-1	0	0	1	1	1	1	-1	-1	0	0	0	0
χ_{13}	$+$	120	120	0	-8	-8	3	3	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	-1	-1	-1	-1
χ_{14}	$+$	144	144	4	0	0	0	0	-3	-3	0	0	-1	-1	1	0	0	0	0	0	0	-1	-1	1	1	1	1
χ_{15}	$+$	176	176	-4	0	0	-4	-4	-1	-1	0	0	1	1	-1	0	0	0	0	0	0	1	1	0	0	0	0
χ_{16}	$\circ$	10	-10	0	-2	2	1	-1	-2	2	0	0	0	0	0	1	-1	a_2	$\overline{a_2}$	a_2	$\overline{a_2}$	0	0	-1	1	-1	1
χ_{17}	$\circ$	10	-10	0	-2	2	1	-1	-2	2	0	0	0	0	0	1	-1	$\overline{a_2}$	a_2	$\overline{a_2}$	a_2	0	0	-1	1	-1	1
χ_{18}	$+$	12	-12	0	4	-4	3	-3	0	0	0	0	2	-2	0	1	-1	0	0	0	0	0	0	1	-1	1	-1
χ_{19}	$-$	32	-32	0	0	0	-4	4	2	-2	0	0	2	-2	0	0	0	0	0	0	0	0	0	-1	1	-1	1
χ_{20}	$\circ$	44	-44	0	4	-4	-1	1	2	-2	0	0	-1	1	0	1	-1	0	0	0	0	a_5	$\overline{a_5}$	0	0	0	0
χ_{21}	$\circ$	44	-44	0	4	-4	-1	1	2	-2	0	0	-1	1	0	1	-1	0	0	0	0	$\overline{a_5}$	a_5	0	0	0	0
χ_{22}	$\circ$	110	-110	0	-6	6	2	-2	2	-2	0	0	0	0	0	0	0	a_2	$\overline{a_2}$	$\overline{a_2}$	a_2	0	0	0	0	0	0
χ_{23}	$\circ$	110	-110	0	-6	6	2	-2	2	-2	0	0	0	0	0	0	0	$\overline{a_2}$	a_2	a_2	$\overline{a_2}$	0	0	0	0	0	0
χ_{24}	$+$	120	-120	0	8	-8	3	-3	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0	-1	1	-1	1
χ_{25}	$\circ$	160	-160	0	0	0	-2	2	-2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	$-b_{11}$	b_{11}	$-\overline{b_{11}}$	$\overline{b_{11}}$
χ_{26}	$\circ$	160	-160	0	0	0	-2	2	-2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	$-\overline{b_{11}}$	$\overline{b_{11}}$	$-b_{11}$	b_{11}

Table 10: Character table of $G^{(4)} \simeq 2.AGL_3(2)$

$[g]$	FS	1A	2A	2B	4A	4B	2C	3A	6A	6B	6C	8A	4C	7A	14A	7B	14B
$[g^2]$		1A	1A	1A	2A	2B	1A	3A	3A	3A	3A	4A	2C	7A	7A	7B	7B
$[g^3]$		1A	2A	2B	4A	4B	2C	1A	2A	2B	2B	8A	4C	7B	14B	7A	14A
$[g^7]$		1A	2A	2B	4A	4B	2C	3A	6A	6B	6C	8A	4C	1A	2A	1A	2A
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	○	3	3	3	-1	-1	-1	0	0	0	0	1	1	b_7	b_7	$\overline{b_7}$	$\overline{b_7}$
χ_3	○	3	3	3	-1	-1	-1	0	0	0	0	1	1	$\overline{b_7}$	$\overline{b_7}$	b_7	b_7
χ_4	+	6	6	6	2	2	2	0	0	0	0	0	0	-1	-1	-1	-1
χ_5	+	7	7	7	-1	-1	-1	1	1	1	1	-1	-1	0	0	0	0
χ_6	+	8	8	8	0	0	0	-1	-1	-1	-1	0	0	1	1	1	1
χ_7	+	7	7	-1	3	-1	-1	1	1	-1	-1	1	-1	0	0	0	0
χ_8	+	7	7	-1	-1	-1	3	1	1	-1	-1	-1	1	0	0	0	0
χ_9	+	14	14	-2	2	-2	2	-1	-1	1	1	0	0	0	0	0	0
χ_{10}	+	21	21	-3	1	1	-3	0	0	0	0	-1	1	0	0	0	0
χ_{11}	+	21	21	-3	-3	1	1	0	0	0	0	1	-1	0	0	0	0
χ_{12}	+	8	-8	0	0	0	0	2	-2	0	0	0	0	1	-1	1	-1
χ_{13}	○	8	-8	0	0	0	0	-1	1	a_3	$\overline{a_3}$	0	0	1	-1	1	-1
χ_{14}	○	8	-8	0	0	0	0	-1	1	$\overline{a_3}$	a_3	0	0	1	-1	1	-1
χ_{15}	○	24	-24	0	0	0	0	0	0	0	0	0	0	$\overline{b_7}$	$-\overline{b_7}$	b_7	$-b_7$
χ_{16}	○	24	-24	0	0	0	0	0	0	0	0	0	0	b_7	$-b_7$	$\overline{b_7}$	$-\overline{b_7}$

Table 11: Character table of $G^{(5)} \simeq GL_2(5)/2$

$[g]$	FS	1A	2A	2B	2C	3A	6A	5A	10A	4A	4B	4C	4D	12A	12B
$[g^2]$		1A	1A	1A	1A	3A	3A	5A	5A	2A	2A	2C	2C	6A	6A
$[g^3]$		1A	2A	2B	2C	1A	2A	5A	10A	4B	4A	4D	4C	4B	4A
$[g^5]$		1A	2A	2B	2C	3A	6A	1A	2A	4A	4B	4C	4D	12A	12B
χ_1	+	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	+	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1
χ_3	+	4	4	0	0	1	1	-1	-1	2	2	0	0	-1	-1
χ_4	+	4	4	0	0	1	1	-1	-1	-2	-2	0	0	1	1
χ_5	+	5	5	1	1	-1	-1	0	0	1	1	-1	-1	1	1
χ_6	+	5	5	1	1	-1	-1	0	0	-1	-1	1	1	-1	-1
χ_7	+	6	6	-2	-2	0	0	1	1	0	0	0	0	0	0
χ_8	○	1	-1	1	-1	1	-1	1	-1	a_1	$-a_1$	a_1	$-a_1$	a_1	$-a_1$
χ_9	○	1	-1	1	-1	1	-1	1	-1	$-a_1$	a_1	$-a_1$	a_1	$-a_1$	a_1
χ_{10}	○	4	-4	0	0	1	-1	-1	1	$2a_1$	$-2a_1$	0	0	$-a_1$	a_1
χ_{11}	○	4	-4	0	0	1	-1	-1	1	$-2a_1$	$2a_1$	0	0	a_1	$-a_1$
χ_{12}	○	5	-5	1	-1	-1	1	0	0	a_1	$-a_1$	$-a_1$	a_1	a_1	$-a_1$
χ_{13}	○	5	-5	1	-1	-1	1	0	0	$-a_1$	a_1	a_1	$-a_1$	$-a_1$	a_1
χ_{14}	+	6	-6	-2	2	0	0	1	-1	0	0	0	0	0	0

Table 12: Character table of $G^{(7)} \simeq SL_2(3)$

$[g]$	FS	1A	2A	4A	3A	6A	3B	6B
$[g^2]$		1A	1A	2A	3B	3A	3A	3B
$[g^3]$		1A	2A	4A	1A	2A	1A	2A
χ_1	$+$	1	1	1	1	1	1	1
χ_2	$\circ$	1	1	1	b_3	$\overline{b_3}$	$\overline{b_3}$	b_3
χ_3	$\circ$	1	1	1	$\overline{b_3}$	b_3	b_3	$\overline{b_3}$
χ_4	$+$	3	3	-1	0	0	0	0
χ_5	$-$	2	-2	0	-1	1	-1	1
χ_6	$\circ$	2	-2	0	$-\overline{b_3}$	b_3	$-b_3$	$\overline{b_3}$
χ_7	$\circ$	2	-2	0	$-b_3$	$\overline{b_3}$	$-\overline{b_3}$	b_3

Table 13: Character table of $G^{(13)} \simeq 4$

$[g]$	FS	1A	2A	4A	4B
$[g^2]$		1A	1A	2A	2A
χ_1	$+$	1	1	1	1
χ_2	$+$	1	1	-1	-1
χ_3	$\circ$	1	-1	a_1	$\overline{a_1}$
χ_4	$\circ$	1	-1	$\overline{a_1}$	a_1

B.2 Euler Characters

The tables in this section describe the Frame shapes $\Pi_g^{(\ell)}$ and twisted Euler characters $\chi_g^{(\ell)}$ attached to each group $G^{(\ell)}$ via the signed permutation representations given in §3.4. The rows labelled $\bar{\Pi}_g^{(\ell)}$ and $\bar{\chi}_g^{(\ell)}$ describe the corresponding data for the (unsigned) permutation representations. According to the discussion of §4.8 the Frame shapes $\Pi_g^{(\ell)}$ and $\bar{\Pi}_g^{(\ell)}$ (or even just the $\Pi_g^{(\ell)}$) can be used to define symbols $n_g|h_g$ which encode the automorphy of the vector-valued mock modular form $H_g^{(\ell)}$; these symbols are given in the rows labelled Γ_g. We write n_g here as a shorthand for $n_g|1$.

Table 14: Twisted Euler characters and Frame shapes at $\ell = 2$

$[g]$	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B
Γ_g	1	2	2|2	3	3|3	4|2	4	4|4	5	6	6|6
$\chi_g^{(2)}$	24	8	0	6	0	0	4	0	4	2	0
$\Pi_g^{(2)}$	1^{24}	$1^8 2^8$	2^{12}	$1^6 3^6$	3^8	$2^4 4^4$	$1^4 2^2 4^4$	4^6	$1^4 5^4$	$1^2 2^2 3^2 6^2$	6^4
$[g]$	7AB	8A	10A	11A	12A	12B	14AB	15AB	21AB	23AB	
Γ_g	7	8	10|2	11	12|2	12|12	14	15	21|3	23	
$\chi_g^{(2)}$	3	2	0	2	0	0	1	1	0	1	
$\Pi_g^{(2)}$	$1^3 7^3$	$1^2 2^1 4^1 8^2$	$2^2 10^2$	$1^2 11^2$	$2^1 4^1 6^1 12^1$	12^2	$1^1 2^1 7^1 14^1$	$1^1 3^1 5^1 15^1$	$3^1 21^1$	$1^1 23^1$	

We have $\chi_g^{(2)} = \chi_1(g) + \chi_2(g)$ in the notation of Table 8.

Table 15: Twisted Euler characters and Frame shapes at $\ell = 3$

$[g]$	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8AB	8CD	20AB	11AB	22AB
Γ_g	1	$1\mid4$	$2\mid8$	2	$2\mid2$	3	$3\mid4$	$3\mid3$	$3\mid12$	$4\mid2$	4	5	$5\mid4$	$6\mid24$	6	$6\mid2$	$8\mid4$	8	$10\mid8$	11	$11\mid4$
$\bar{\chi}_g^{(3)}$	12	12	0	4	4	3	3	0	0	0	4	2	2	0	1	1	0	2	0	1	1
$\chi_g^{(3)}$	12	−12	0	4	−4	3	−3	0	0	0	0	2	−2	0	1	−1	0	0	0	1	−1
$\bar{\Pi}_g^{(3)}$	1^{12}	1^{12}	2^6	1^42^4	1^42^4	1^33^3	1^33^3	3^4	3^4	2^24^2	1^44^2	1^25^2	1^25^2	6^2	$1^12^13^16^1$	$1^12^13^16^1$	4^18^1	$1^22^18^1$	2^110^1	1^111^1	1^111^1
$\Pi_g^{(3)}$	1^{12}	$\frac{2^{12}}{1^{12}}$	$\frac{4^6}{2^6}$	1^42^4	$\frac{2^8}{1^4}$	1^33^3	$\frac{2^36^3}{1^33^3}$	3^4	$\frac{6^4}{3^4}$	2^24^2	2^24^2	1^25^2	$\frac{2^210^2}{1^25^2}$	$\frac{12^2}{6^2}$	$1^12^13^16^1$	$\frac{2^26^2}{1^13^1}$	4^18^1	4^18^1	$\frac{4^120^1}{2^110^1}$	1^111^1	$\frac{2^122^1}{1^111^1}$

We have $\chi_g^{(3)} = \chi_{18}(g)$ and $\bar{\chi}_g^{(3)} = \chi_1(g) + \chi_2(g)$ in the notation of Table 9.

Table 16: Twisted Euler characters and Frame shapes at $\ell = 4$

$[g]$	1A	2A	2B	4A	4B	2C	3A	6A	6BC	8A	4C	7AB	14AB
Γ_g	1	$1\mid2$	$2\mid2$	$2\mid4$	$4\mid4$	2	3	$3\mid2$	$6\mid2$	$4\mid8$	4	7	$7\mid2$
$\bar{\chi}_g^{(4)}$	8	8	0	0	0	4	2	2	0	0	2	1	1
$\chi_g^{(4)}$	8	−8	0	0	0	0	2	−2	0	0	0	1	−1
$\bar{\Pi}_g^{(4)}$	1^8	1^8	2^4	2^4	4^2	1^42^2	1^23^2	1^23^2	2^16^1	4^2	$1^22^14^1$	1^17^1	1^17^1
$\Pi_g^{(4)}$	1^8	$\frac{2^8}{1^8}$	2^4	$\frac{4^4}{2^4}$	4^2	2^4	1^23^2	$\frac{2^26^2}{1^23^2}$	2^16^1	$\frac{8^2}{4^2}$	4^2	1^17^1	$\frac{2^114^1}{1^17^1}$

We have $\chi_g^{(4)} = \chi_{12}$ and $\bar{\chi}_g^{(4)} = \chi_1(g) + \chi_8(g)$ in the notation of Table 10.

Table 17: Twisted Euler characters and Frame shapes at $\ell = 5$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
Γ_g	1	1|4	2|2	2	3|3	3|12	5	5|4	2|8	4	6|24
$\bar{\chi}_g^{(5)}$	6	6	2	2	0	0	1	1	0	2	0
$\chi_g^{(5)}$	6	-6	-2	2	0	0	1	-1	0	0	0
$\bar{\Pi}_g^{(5)}$	1^6	1^6	$1^2 2^2$	$1^2 2^2$	3^2	3^2	$1^1 5^1$	$1^1 5^1$	2^3	$1^2 4^1$	6^1
$\Pi_g^{(5)}$	1^6	$\frac{2^6}{1^6}$	$\frac{2^4}{1^2}$	$1^2 2^2$	3^2	$\frac{6^2}{3^2}$	$1^1 5^1$	$\frac{2^1 10^1}{1^1 5^1}$	$\frac{4^3}{2^3}$	$2^1 4^1$	$\frac{12^1}{6^1}$

We have $\chi_g^{(5)} = \chi_{14}(g)$ and $\bar{\chi}_g^{(5)} = \chi_1(g) + \chi_6(g)$ in the notation of Table 11.

Table 18: Twisted Euler characters and Frame shapes at $\ell = 7$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
$\bar{\chi}_g^{(7)}$	4	4	0	1	1
$\chi_g^{(7)}$	4	-4	0	1	-1
$\bar{\Pi}_g^{(7)}$	1^4	1^4	2^2	$1^1 3^1$	$1^1 3^1$
$\Pi_g^{(7)}$	1^4	$\frac{2^4}{1^4}$	$\frac{4^2}{2^2}$	$1^1 3^1$	$\frac{2^1 6^1}{1^1 3^1}$

We have $\chi_g^{(7)} = \chi_6(g) + \chi_7(g)$ in the notation of Table 12.

Table 19: Twisted Euler characters and Frame shapes at $\ell = 13$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
$\bar{\chi}_g^{(13)}$	2	2	0
$\chi_g^{(13)}$	2	-2	0
$\bar{\Pi}_g^{(13)}$	1^2	1^2	2^1
$\Pi_g^{(13)}$	1^2	$\frac{2^2}{1^2}$	$\frac{4^1}{2^1}$

We have $\chi_g^{(13)} = \chi_3(g) + \chi_4(g)$ in the notation of Table 13.

C Coefficients

In this section we furnish tables of Fourier coefficients of small degree for the vector-valued mock modular forms $H^{(\ell)}_g$ that we attach to the conjugacy classes of the groups $G^{(\ell)}$ for $\ell \in \Lambda$. For each ℓ and $0 < r < \ell$ we give a table that displays the coefficients of $H^{(\ell)}_{g,r}$ for (g ranging over a set of representatives for) each conjugacy class $[g]$ in $G^{(\ell)}$. The first row of each table labels the conjugacy classes, and the first column labels exponents of q (or rather $q^{1/4\ell}$), so that for the table captioned $H^{(\ell)}_{g,r}$ (for some $\ell \in \Lambda$ and $0 < r < \ell$) the entry in the row labelled d and the column labelled nZ is the coefficient of $q^{d/4\ell}$ in the Fourier expansion of $H^{(\ell)}_{g,r}$ for $[g] = nZ$. Occasionally the functions $H^{(\ell)}_g$ and $H^{(\ell)}_{g'}$ coincide for non-conjugate g and g' and when this happens we condense information into a single column, writing $7AB$ in Table 20, for example, to indicate that the entries in that column are Fourier coefficients for both $H^{(2)}_{7A}$ and $H^{(2)}_{7B}$.

C.1 Lambency Two

Table 20: McKay–Thompson series $H^{(2)}_{g,1}$

$[g]$	1A	2A	2B	3A	3B	4A	4B	4C	5A	6A	6B	7AB	8A	10A	11A	12A	12B	14AB	15AB	21AB	23AB
Γ_g	1	2	2|2	3	3|3	4|2	4	4|4	5	6	6|6	7	8	10|2	11	12|2	12|12	14	15	21|3	23
-1	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
7	90	-6	10	0	6	-6	2	2	0	0	-2	-1	-2	0	2	0	2	1	0	-1	-2
15	462	14	-18	-6	0	-2	-2	6	2	2	0	0	-2	2	0	-2	0	0	-1	0	2
23	1540	-28	20	10	-14	4	-4	-4	0	2	2	0	0	0	0	-2	2	0	0	0	-1
31	4554	42	-38	0	12	-6	2	-6	-6	0	4	4	-2	2	0	0	0	0	0	-2	0
39	11592	-56	72	-18	0	-8	8	0	2	-2	0	0	0	2	-2	-2	0	0	2	0	0
47	27830	86	-90	20	-16	6	-2	6	0	-4	0	-2	2	0	0	0	0	2	0	-2	0
55	61686	-138	118	0	30	6	-10	-2	6	0	-2	2	-2	-2	-2	0	-2	2	0	2	0
63	131100	188	-180	-30	0	-4	4	-12	0	2	0	-3	0	0	2	2	0	-1	0	0	0
71	265650	-238	258	42	-42	-14	10	10	-10	2	6	0	-2	-2	0	-2	-2	0	2	0	0
79	521136	336	-352	0	42	0	-8	16	6	0	2	0	-4	-2	0	0	-2	0	0	0	2
87	988770	-478	450	-60	0	18	-14	-6	0	-4	0	6	2	0	2	0	0	-2	0	0	0
95	1830248	616	-600	62	-70	-8	8	-16	8	-2	-6	0	0	0	2	-2	2	0	2	0	0
103	3303630	-786	830	0	84	-18	22	6	0	0	-4	-6	2	0	0	0	0	-2	0	0	2
111	5844762	1050	-1062	-90	0	10	-6	18	-18	6	0	0	2	-2	0	-2	0	0	0	0	2
119	10139734	-1386	1334	118	-110	22	-26	-10	4	6	2	-4	-2	4	0	-2	2	0	-2	2	0
127	17301060	1764	-1740	0	126	-12	12	-28	0	0	6	0	0	0	-4	0	2	0	0	0	0
135	29051484	-2212	2268	-156	0	-36	28	12	14	-4	0	0	-4	-2	0	0	0	0	-1	0	0
143	48106430	2814	-2850	170	-166	14	-18	38	0	-6	-6	8	-2	0	-2	2	2	0	0	2	-2
151	78599556	-3612	3540	0	210	36	-36	-20	-24	0	-6	0	0	0	2	0	-2	0	0	0	0
159	126894174	4510	-4482	-228	0	-18	14	-42	14	4	0	-6	-2	-2	0	0	0	2	2	0	0
167	202537080	-5544	5640	270	-282	-40	48	16	0	6	6	4	4	0	-2	2	-2	0	0	-2	0
175	319927608	6936	-6968	0	300	24	-16	48	18	0	4	-7	4	2	0	0	0	-1	0	-1	0
183	500376870	-8666	8550	-360	0	54	-58	-18	0	-8	0	0	-2	0	4	0	0	0	0	0	2
191	775492564	10612	-10556	400	-392	-28	28	-60	-36	-8	-8	0	0	4	0	-4	0	0	0	0	0
199	1191453912	-12936	13064	0	462	-72	64	32	12	0	-10	12	-4	4	0	0	2	0	0	0	0
207	1815754710	15862	-15930	-510	0	22	-34	78	0	10	0	0	-6	0	0	-2	0	0	0	0	-1
215	2745870180	-19420	19268	600	-600	84	-76	-36	30	8	8	-10	4	-2	-2	0	0	-2	0	2	0
223	4122417420	23532	-23460	0	660	-36	36	-84	0	0	12	2	0	0	0	0	0	-2	0	2	0
231	6146311620	-28348	28548	-762	0	-92	100	36	-50	-10	0	-6	4	-2	-2	-2	0	2	-2	0	0
239	9104078592	34272	-34352	828	-840	48	-40	96	22	-12	-8	0	4	-2	4	0	0	0	-2	0	0
247	13401053820	-41412	41180	0	966	108	-116	-44	0	0	-10	0	-4	0	0	0	-2	0	0	0	-2

C.2 Lambency Three

Table 21: McKay–Thompson series $H^{(3)}_{g,1}$

$[g]$	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8AB	8CD	20AB	11AB	22AB
Γ_g	1	1|4	2|8	2	2|2	3	3|4	3|3	3|12	4|2	4	5	5|4	6|24	6	6|2	8|4	8	10|8	11	11|4
-1	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
11	32	32	8	0	0	-4	-4	2	2	0	0	2	2	2	0	0	0	0	-2	-1	-1
23	110	110	-10	-2	-2	2	2	2	2	6	-2	0	0	2	-2	-2	-2	2	0	0	0
35	288	288	8	0	0	0	0	-6	-6	0	0	-2	-2	2	0	0	0	0	-2	2	2
47	660	660	-20	4	4	-6	-6	6	6	-4	4	0	0	-2	-2	-2	4	0	0	0	0
59	1408	1408	32	0	0	4	4	4	4	0	0	-2	-2	-4	0	0	0	0	2	0	0
71	2794	2794	-30	-6	-6	4	4	-8	-8	2	-6	4	4	0	0	0	2	-2	0	0	0
83	5280	5280	40	0	0	-12	-12	6	6	0	0	0	0	-2	0	0	0	0	0	0	0
95	9638	9638	-58	6	6	8	8	2	2	-10	6	-2	-2	2	0	0	-2	-2	2	2	2
107	16960	16960	80	0	0	4	4	-14	-14	0	0	0	0	2	0	0	0	0	0	-2	-2
119	29018	29018	-102	-6	-6	-16	-16	8	8	10	-6	-2	-2	0	0	0	2	2	-2	0	0
131	48576	48576	112	0	0	12	12	6	6	0	0	6	6	-2	0	0	0	0	2	0	0
143	79530	79530	-150	10	10	6	6	-24	-24	-6	10	0	0	0	-2	-2	2	2	0	0	0
155	127776	127776	200	0	0	-24	-24	18	18	0	0	-4	-4	2	0	0	0	0	0	0	0
167	202050	202050	-230	-14	-14	18	18	12	12	10	-14	0	0	4	-2	-2	2	-2	0	2	2
179	314688	314688	272	0	0	12	12	-30	-30	0	0	-2	-2	2	0	0	0	0	2	0	0
191	483516	483516	-348	12	12	-36	-36	24	24	-12	12	6	6	0	0	0	-4	0	2	0	0
203	733920	733920	440	0	0	24	24	12	12	0	0	0	0	-4	0	0	0	0	0	0	0
215	1101364	1101364	-508	-12	-12	16	16	-44	-44	20	-12	-6	-6	-4	0	0	-4	4	2	0	0
227	1635680	1635680	600	0	0	-52	-52	32	32	0	0	0	0	0	0	0	0	0	0	2	2
239	2406116	2406116	-740	20	20	38	38	20	20	-20	20	-4	-4	4	2	2	4	0	0	-2	-2
251	3507680	3507680	888	0	0	20	20	-64	-64	0	0	10	10	0	0	0	0	0	-2	0	0
263	5071000	5071000	-1040	-24	-24	-68	-68	46	46	16	-24	0	0	-2	0	0	0	-4	0	0	0
275	7274464	7274464	1208	0	0	52	52	28	28	0	0	-6	-6	-4	0	0	0	0	-2	-1	-1
287	10359030	10359030	-1450	22	22	30	30	-84	-84	-26	22	0	0	-4	-2	-2	-2	-2	0	0	0
299	14650176	14650176	1744	0	0	-96	-96	60	60	0	0	-4	-4	4	0	0	0	0	4	2	2
311	20585334	20585334	-2018	-26	-26	66	66	36	36	30	-26	14	14	4	-2	-2	-2	2	2	0	0
323	28747840	28747840	2320	0	0	40	40	-116	-116	0	0	0	0	4	0	0	0	0	0	0	0
335	39914402	39914402	-2750	34	34	-130	-130	86	86	-30	34	-8	-8	-2	-2	-2	2	2	0	0	0
347	55114400	55114400	3240	0	0	92	92	50	50	0	0	0	0	-6	0	0	0	0	0	0	0
359	75704904	75704904	-3712	-40	-40	54	54	-156	-156	32	-40	-6	-6	-4	2	2	0	-4	-2	0	0

Table 22: McKay–Thompson series $H^{(3)}_{g,2}$

$[g]$	1A	2A	4A	2B	2C	3A	6A	3B	6B	4B	4C	5A	10A	12A	6C	6D	8AB	8CD	20AB	11AB	22AB
Γ_g	1	1|4	2|8	2	2|2	3	3|4	3|3	3|12	4|2	4	5	5|4	6|24	6	6|2	8|4	8	10|8	11	11|4
8	20	-20	0	-4	4	2	-2	-4	4	0	0	0	0	0	2	-2	0	0	0	-2	2
20	88	-88	0	8	-8	-2	2	4	-4	0	0	-2	2	0	2	-2	0	0	0	0	0
32	220	-220	0	-12	12	4	-4	4	-4	0	0	0	0	0	0	0	0	0	0	0	0
44	560	-560	0	16	-16	2	-2	-4	4	0	0	0	0	0	-2	2	0	0	0	-1	1
56	1144	-1144	0	-24	24	-8	8	4	-4	0	0	4	-4	0	0	0	0	0	0	0	0
68	2400	-2400	0	32	-32	6	-6	0	0	0	0	0	0	0	2	-2	0	0	0	2	-2
80	4488	-4488	0	-40	40	6	-6	-12	12	0	0	-2	2	0	2	-2	0	0	0	0	0
92	8360	-8360	0	56	-56	-10	10	8	-8	0	0	0	0	0	2	-2	0	0	0	0	0
104	14696	-14696	0	-72	72	8	-8	8	-8	0	0	-4	4	0	0	0	0	0	0	0	0
116	25544	-25544	0	88	-88	2	-2	-16	16	0	0	4	-4	0	-2	2	0	0	0	2	-2
128	42660	-42660	0	-116	116	-18	18	12	-12	0	0	0	0	0	-2	2	0	0	0	2	-2
140	70576	-70576	0	144	-144	16	-16	4	-4	0	0	-4	4	0	0	0	0	0	0	0	0
152	113520	-113520	0	-176	176	12	-12	-24	24	0	0	0	0	0	4	-4	0	0	0	0	0
164	180640	-180640	0	224	-224	-26	26	16	-16	0	0	0	0	0	2	-2	0	0	0	-2	2
176	281808	-281808	0	-272	272	18	-18	12	-12	0	0	8	-8	0	-2	2	0	0	0	-1	1
188	435160	-435160	0	328	-328	10	-10	-32	32	0	0	0	0	0	-2	2	0	0	0	0	0
200	661476	-661476	0	-404	404	-42	42	24	-24	0	0	-4	4	0	-2	2	0	0	0	2	-2
212	996600	-996600	0	488	-488	30	-30	12	-12	0	0	0	0	0	2	-2	0	0	0	0	0
224	1482536	-1482536	0	-584	584	20	-20	-52	52	0	0	-4	4	0	4	-4	0	0	0	0	0
236	2187328	-2187328	0	704	-704	-50	50	40	-40	0	0	8	-8	0	2	-2	0	0	0	0	0
248	3193960	-3193960	0	-840	840	40	-40	28	-28	0	0	0	0	0	0	0	0	0	0	0	0
260	4629152	-4629152	0	992	-992	20	-20	-64	64	0	0	-8	8	0	-4	4	0	0	0	0	0
272	6650400	-6650400	0	-1184	1184	-78	78	48	-48	0	0	0	0	0	-2	2	0	0	0	-2	2
284	9490536	-9490536	0	1400	-1400	54	-54	24	-24	0	0	-4	4	0	2	-2	0	0	0	0	0
296	13441032	-13441032	0	-1640	1640	36	-36	-96	96	0	0	12	-12	0	4	-4	0	0	0	0	0
308	18920240	-18920240	0	1936	-1936	-100	100	68	-68	0	0	0	0	0	4	-4	0	0	0	-2	2
320	26457464	-26457464	0	-2264	2264	74	-74	44	-44	0	0	-6	6	0	-2	2	0	0	0	0	0
332	36792560	-36792560	0	2640	-2640	38	-38	-124	124	0	0	0	0	0	-6	6	0	0	0	2	-2
344	50865232	-50865232	0	-3088	3088	-140	140	88	-88	0	0	-8	8	0	-4	4	0	0	0	0	0
356	69966336	-69966336	0	3584	-3584	102	-102	48	-48	0	0	16	-16	0	2	-2	0	0	0	0	0

C.3 Lambency Four

Table 23: McKay–Thompson series $H^{(4)}_{g,1}$

$[g]$	1A	2A	2B	4A	4B	2C	3A	6A	6BC	8A	4C	7AB	14AB
Γ_g	1	$1\vert 2$	$2\vert 2$	$2\vert 4$	$4\vert 4$	2	3	$3\vert 2$	$6\vert 2$	$4\vert 8$	4	7	$7\vert 2$
-1	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
15	14	14	-2	6	-2	-2	2	2	-2	2	-2	0	0
31	42	42	-6	-6	2	2	0	0	0	2	-2	0	0
47	86	86	6	6	-2	-2	-4	-4	0	-2	2	2	2
63	188	188	-4	-12	-4	4	2	2	2	0	0	-1	-1
79	336	336	0	16	0	-8	0	0	0	0	-4	0	0
95	616	616	-8	-16	0	8	-2	-2	-2	4	0	0	0
111	1050	1050	10	18	2	-6	6	6	-2	-2	2	0	0
127	1764	1764	-12	-28	-4	12	0	0	0	-4	0	0	0
143	2814	2814	14	38	-2	-18	-6	-6	2	2	-2	0	0
159	4510	4510	-18	-42	6	14	4	4	0	2	-2	2	2
175	6936	6936	24	48	0	-16	0	0	0	-4	4	-1	-1
191	10612	10612	-28	-60	-4	28	-8	-8	-4	-4	0	0	0
207	15862	15862	22	78	-2	-34	10	10	-2	2	-6	0	0
223	23532	23532	-36	-84	4	36	0	0	0	4	0	-2	-2
239	34272	34272	48	96	0	-40	-12	-12	0	0	4	0	0
255	49618	49618	-46	-126	-6	50	10	10	2	-6	2	2	2
271	70758	70758	54	150	-2	-66	0	0	0	6	-6	2	2
287	100310	100310	-74	-170	6	70	-10	-10	-2	6	-2	0	0
303	140616	140616	88	192	0	-72	18	18	-2	-4	8	0	0
319	195888	195888	-96	-232	-8	96	0	0	0	-4	0	0	0
335	270296	270296	104	272	0	-120	-22	-22	2	4	-8	-2	-2
351	371070	371070	-130	-306	6	126	18	18	2	6	-2	0	0
367	505260	505260	156	348	4	-140	0	0	0	-4	8	0	0
383	684518	684518	-170	-410	-10	174	-22	-22	-2	-10	2	2	2
399	921142	921142	182	486	-2	-202	28	28	-4	6	-10	-2	-2
415	1233708	1233708	-228	-540	12	220	0	0	0	8	-4	0	0
431	1642592	1642592	272	608	0	-248	-34	-34	2	-8	12	0	0
447	2177684	2177684	-284	-708	-12	292	32	32	4	-8	4	-2	-2
463	2871918	2871918	318	814	-2	-346	0	0	0	6	-14	0	0
479	3772468	3772468	-380	-908	12	380	-38	-38	-2	12	0	0	0
495	4932580	4932580	436	1020	4	-412	46	46	-2	-8	12	2	2
511	6425466	6425466	-486	-1174	-14	490	0	0	0	-14	2	-2	-2
527	8335418	8335418	538	1338	-6	-566	-52	-52	4	10	-14	0	0
543	10776290	10776290	-622	-1494	18	610	50	50	2	14	-6	0	0
559	13879290	13879290	714	1666	2	-678	0	0	0	-10	18	-2	-2
575	17818766	17818766	-786	-1898	-18	790	-58	-58	-6	-14	2	0	0
591	22798188	22798188	860	2148	-4	-900	72	72	-4	8	-20	0	0

Table 24: McKay–Thompson series $H^{(4)}_{g,2}$

$[g]$	1A	2A	2B	4A	4B	2C	3A	6A	6BC	8A	4C	7AB	14AB
Γ_g	1	1|2	2|2	2|4	4|4	2	3	3|2	6|2	4|8	4	7	7|2
12	16	-16	0	0	0	0	-2	2	0	0	0	2	-2
28	48	-48	0	0	0	0	0	0	0	0	0	-1	1
44	112	-112	0	0	0	0	4	-4	0	0	0	0	0
60	224	-224	0	0	0	0	-4	4	0	0	0	0	0
76	432	-432	0	0	0	0	0	0	0	0	0	-2	2
92	784	-784	0	0	0	0	4	-4	0	0	0	0	0
108	1344	-1344	0	0	0	0	-6	6	0	0	0	0	0
124	2256	-2256	0	0	0	0	0	0	0	0	0	2	-2
140	3680	-3680	0	0	0	0	8	-8	0	0	0	-2	2
156	5824	-5824	0	0	0	0	-8	8	0	0	0	0	0
172	9072	-9072	0	0	0	0	0	0	0	0	0	0	0
188	13872	-13872	0	0	0	0	12	-12	0	0	0	-2	2
204	20832	-20832	0	0	0	0	-12	12	0	0	0	0	0
220	30912	-30912	0	0	0	0	0	0	0	0	0	0	0
236	45264	-45264	0	0	0	0	12	-12	0	0	0	2	-2
252	65456	-65456	0	0	0	0	-16	16	0	0	0	-1	1
268	93744	-93744	0	0	0	0	0	0	0	0	0	0	0
284	132944	-132944	0	0	0	0	20	-20	0	0	0	0	0
300	186800	-186800	0	0	0	0	-22	22	0	0	0	-2	2
316	260400	-260400	0	0	0	0	0	0	0	0	0	0	0
332	360208	-360208	0	0	0	0	28	-28	0	0	0	2	-2
348	494624	-494624	0	0	0	0	-28	28	0	0	0	4	-4
364	674784	-674784	0	0	0	0	0	0	0	0	0	-2	2
380	914816	-914816	0	0	0	0	32	-32	0	0	0	0	0
396	1232784	-1232784	0	0	0	0	-36	36	0	0	0	0	0
412	1652208	-1652208	0	0	0	0	0	0	0	0	0	-2	2
428	2202704	-2202704	0	0	0	0	44	-44	0	0	0	0	0
444	2921856	-2921856	0	0	0	0	-48	48	0	0	0	0	0
460	3857760	-3857760	0	0	0	0	0	0	0	0	0	4	-4
476	5070560	-5070560	0	0	0	0	56	-56	0	0	0	-2	2
492	6636000	-6636000	0	0	0	0	-60	60	0	0	0	0	0
508	8649648	-8649648	0	0	0	0	0	0	0	0	0	0	0
524	11230448	-11230448	0	0	0	0	68	-68	0	0	0	-2	2
540	14526848	-14526848	0	0	0	0	-76	76	0	0	0	0	0
556	18724176	-18724176	0	0	0	0	0	0	0	0	0	2	-2
572	24051808	-24051808	0	0	0	0	88	-88	0	0	0	4	-4
588	30793712	-30793712	0	0	0	0	-94	94	0	0	0	-2	2
604	39301584	-39301584	0	0	0	0	0	0	0	0	0	0	0

Table 25: McKay–Thompson series $H_{g,3}^{(4)}$

$[g]$	1A	2A	2B	4A	4B	2C	3A	6A	6BC	8A	4C	7AB	14AB
Γ_g	1	1|2	2|2	2|4	4|4	2	3	3|2	6|2	4|8	4	7	7|2
7	6	6	6	-2	-2	-2	0	0	0	2	2	-1	-1
23	28	28	-4	4	-4	4	-2	-2	2	0	0	0	0
39	56	56	8	0	0	-8	2	2	2	-4	0	0	0
55	138	138	-6	2	2	10	0	0	0	-2	2	-2	-2
71	238	238	14	-10	-2	-10	-2	-2	2	2	2	0	0
87	478	478	-18	6	-2	14	4	4	0	2	-2	2	2
103	786	786	18	-6	2	-22	0	0	0	-2	-2	2	2
119	1386	1386	-22	10	2	26	-6	-6	2	2	2	0	0
135	2212	2212	36	-12	-4	-28	4	4	0	4	4	0	0
151	3612	3612	-36	20	-4	36	0	0	0	0	0	0	0
167	5544	5544	40	-16	0	-48	-6	-6	-2	-4	-4	0	0
183	8666	8666	-54	18	2	58	8	8	0	-2	2	0	0
199	12936	12936	72	-32	0	-64	0	0	0	4	4	0	0
215	19420	19420	-84	36	-4	76	-8	-8	0	0	-4	2	2
231	28348	28348	92	-36	4	-100	10	10	2	-4	-4	-2	-2
247	41412	41412	-108	44	4	116	0	0	0	0	4	0	0
263	59178	59178	138	-62	-6	-126	-12	-12	0	6	6	0	0
279	84530	84530	-158	66	-6	154	14	14	-2	2	-2	-2	-2
295	118692	118692	180	-68	4	-188	0	0	0	-8	-4	0	0
311	166320	166320	-208	88	8	216	-18	-18	2	-4	4	0	0
327	230092	230092	252	-108	-4	-244	16	16	0	8	4	2	2
343	317274	317274	-294	122	-6	282	0	0	0	2	-6	-1	-1
359	432964	432964	324	-132	4	-340	-20	-20	0	-8	-8	0	0
375	588966	588966	-378	150	6	390	24	24	0	-2	6	0	0
391	794178	794178	450	-190	-6	-430	0	0	0	10	10	0	0
407	1067220	1067220	-508	220	-12	500	-30	-30	2	0	-4	0	0
423	1423884	1423884	572	-228	4	-588	30	30	2	-12	-8	0	0
439	1893138	1893138	-654	266	10	666	0	0	0	-2	6	2	2
455	2501434	2501434	762	-326	-6	-742	-32	-32	0	10	10	-2	-2
471	3294256	3294256	-864	360	-8	848	40	40	0	4	-8	0	0
487	4314912	4314912	960	-392	8	-984	0	0	0	-12	-12	0	0
503	5633596	5633596	-1092	452	12	1108	-44	-44	0	0	8	-4	-4
519	7320670	7320670	1262	-522	-10	-1234	46	46	2	18	14	0	0
535	9483336	9483336	-1416	592	-16	1400	0	0	0	4	-8	2	2
551	12233330	12233330	1570	-646	10	-1598	-58	-58	-2	-18	-14	4	4
567	15734606	15734606	-1778	726	14	1798	62	62	-2	-6	10	-1	-1
583	20161302	20161302	2022	-850	-10	-1994	0	0	0	18	14	0	0
599	25761288	25761288	-2264	944	-16	2240	-72	-72	3	4	-12	0	0

C.4 Lambency Five

Table 26: McKay–Thompson series $H^{(5)}_{g,1}$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
Γ_g	1	1|4	2|2	2	3|3	3|12	5	5|4	2|8	4	6|24
-1	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
19	8	8	0	0	2	2	-2	-2	4	0	-2
39	18	18	2	2	0	0	-2	-2	-6	2	0
59	40	40	0	0	-2	-2	0	0	4	0	-2
79	70	70	-2	-2	4	4	0	0	-6	-2	0
99	120	120	0	0	0	0	0	0	12	0	0
119	208	208	0	0	-2	-2	-2	-2	-8	0	-2
139	328	328	0	0	4	4	-2	-2	12	0	0
159	510	510	-2	-2	0	0	0	0	-18	-2	0
179	792	792	0	0	-6	-6	2	2	20	0	2
199	1180	1180	4	4	4	4	0	0	-24	4	0
219	1728	1728	0	0	0	0	-2	-2	24	0	0
239	2518	2518	-2	-2	-8	-8	-2	-2	-30	-2	0
259	3600	3600	0	0	6	6	0	0	40	0	-2
279	5082	5082	2	2	0	0	2	2	-42	2	0
299	7120	7120	0	0	-8	-8	0	0	48	0	0
319	9838	9838	-2	-2	10	10	-2	-2	-58	-2	2
339	13488	13488	0	0	0	0	-2	-2	72	0	0
359	18380	18380	4	4	-10	-10	0	0	-80	4	-2
379	24792	24792	0	0	12	12	2	2	84	0	0
399	33210	33210	-6	-6	0	0	0	0	-102	-6	0
419	44248	44248	0	0	-14	-14	-2	-2	116	0	2
439	58538	58538	2	2	14	14	-2	-2	-130	2	2
459	76992	76992	0	0	0	0	2	2	144	0	0
479	100772	100772	-4	-4	-16	-16	2	2	-168	-4	0
499	131160	131160	0	0	18	18	0	0	196	0	-2
519	169896	169896	8	8	0	0	-4	-4	-216	8	0
539	219128	219128	0	0	-22	-22	-2	-2	236	0	2
559	281322	281322	-6	-6	24	24	2	2	-270	-6	0
579	359712	359712	0	0	0	0	2	2	312	0	0
599	458220	458220	4	4	-24	-24	0	0	-336	4	0
619	581416	581416	0	0	28	28	-4	-4	372	0	0
639	735138	735138	-6	-6	0	0	-2	-2	-426	-6	0
659	926472	926472	0	0	-30	-30	2	2	476	0	2
679	1163674	1163674	10	10	34	34	4	4	-526	10	2
699	1457040	1457040	0	0	0	0	0	0	576	0	0
719	1819056	1819056	-8	-8	-42	-42	-4	-4	-644	-8	-2
739	2264376	2264376	0	0	42	42	-4	-4	724	0	-2

Table 27: McKay–Thompson series $H^{(5)}_{g,2}$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
Γ_g	1	1|4	2|2	2	3|3	3|12	5	5|4	2|8	4	6|24
16	10	-10	2	-2	-2	2	0	0	0	0	0
36	30	-30	-2	2	0	0	0	0	0	0	0
56	52	-52	4	-4	4	-4	2	-2	0	0	0
76	108	-108	-4	4	0	0	-2	2	0	0	0
96	180	-180	4	-4	0	0	0	0	0	0	0
116	312	-312	-8	8	0	0	2	-2	0	0	0
136	488	-488	8	-8	-4	4	-2	2	0	0	0
156	792	-792	-8	8	0	0	2	-2	0	0	0
176	1180	-1180	12	-12	4	-4	0	0	0	0	0
196	1810	-1810	-14	14	-2	2	0	0	0	0	0
216	2640	-2640	16	-16	0	0	0	0	0	0	0
236	3868	-3868	-20	20	4	-4	-2	2	0	0	0
256	5502	-5502	22	-22	-6	6	2	-2	0	0	0
276	7848	-7848	-24	24	0	0	-2	2	0	0	0
296	10912	-10912	32	-32	4	-4	2	-2	0	0	0
316	15212	-15212	-36	36	-4	4	2	-2	0	0	0
336	20808	-20808	40	-40	0	0	-2	2	0	0	0
356	28432	-28432	-48	48	4	-4	2	-2	0	0	0
376	38308	-38308	52	-52	-8	8	-2	2	0	0	0
396	51540	-51540	-60	60	0	0	0	0	0	0	0
416	68520	-68520	72	-72	12	-12	0	0	0	0	0
436	90928	-90928	-80	80	-8	8	-2	2	0	0	0
456	119544	-119544	88	-88	0	0	4	-4	0	0	0
476	156728	-156728	-104	104	8	-8	-2	2	0	0	0
496	203940	-203940	116	-116	-12	12	0	0	0	0	0
516	264672	-264672	-128	128	0	0	2	-2	0	0	0
536	341188	-341188	148	-148	16	-16	-2	2	0	0	0
556	438732	-438732	-164	164	-12	12	2	-2	0	0	0
576	560958	-560958	182	-182	0	0	-2	2	0	0	0
596	715312	-715312	-208	208	16	-16	2	-2	0	0	0
616	907720	-907720	232	-232	-20	20	0	0	0	0	0
636	1148928	-1148928	-256	256	0	0	-2	2	0	0	0
656	1447904	-1447904	288	-288	20	-20	4	-4	0	0	0
676	1820226	-1820226	-318	318	-18	18	-4	4	0	0	0
696	2279520	-2279520	352	-352	0	0	0	0	0	0	0
716	2847812	-2847812	-396	396	20	-20	2	-2	0	0	0
736	3545636	-3545636	436	-436	-28	28	-4	4	0	0	0
756	4404384	-4404384	-480	480	0	0	4	-4	0	0	0

Table 28: McKay–Thompson series $H^{(5)}_{g,3}$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
Γ_g	1	1|4	2|2	2	3|3	3|12	5	5|4	2|8	4	6|24
11	8	8	0	0	2	2	-2	-2	-4	0	2
31	22	22	-2	-2	-2	-2	2	2	2	-2	2
51	48	48	0	0	0	0	-2	-2	0	0	0
71	90	90	2	2	0	0	0	0	6	2	0
91	160	160	0	0	-2	-2	0	0	-8	0	-2
111	270	270	-2	-2	0	0	0	0	6	-2	0
131	440	440	0	0	2	2	0	0	-4	0	2
151	700	700	4	4	-2	-2	0	0	8	4	2
171	1080	1080	0	0	0	0	0	0	-12	0	0
191	1620	1620	-4	-4	6	6	0	0	16	-4	-2
211	2408	2408	0	0	-4	-4	-2	-2	-12	0	0
231	3522	3522	2	2	0	0	2	2	18	2	0
251	5048	5048	0	0	2	2	-2	-2	-28	0	2
271	7172	7172	-4	-4	-4	-4	2	2	24	-4	0
291	10080	10080	0	0	0	0	0	0	-24	0	0
311	13998	13998	6	6	6	6	-2	-2	34	6	-2
331	19272	19272	0	0	-6	-6	2	2	-44	0	-2
351	26298	26298	-6	-6	0	0	-2	-2	42	-6	0
371	35600	35600	0	0	8	8	0	0	-48	0	0
391	47862	47862	6	6	-6	-6	2	2	62	6	2
411	63888	63888	0	0	0	0	-2	-2	-72	0	0
431	84722	84722	-6	-6	8	8	2	2	78	-6	0
451	111728	111728	0	0	-10	-10	-2	-2	-80	0	-2
471	146520	146520	8	8	0	0	0	0	96	8	0
491	191080	191080	0	0	10	10	0	0	-124	0	2
511	248008	248008	-8	-8	-14	-14	-2	-2	128	-8	2
531	320424	320424	0	0	0	0	4	4	-132	0	0
551	412088	412088	8	8	14	14	-2	-2	160	8	-2
571	527800	527800	0	0	-14	-14	0	0	-188	0	-2
591	673302	673302	-10	-10	0	0	2	2	198	-10	0
611	855616	855616	0	0	16	16	-4	-4	-216	0	0
631	1083444	1083444	12	12	-18	-18	4	4	248	12	2
651	1367136	1367136	0	0	0	0	-4	-4	-288	0	0
671	1719362	1719362	-14	-14	20	20	2	2	314	-14	-4
691	2155592	2155592	0	0	-22	-22	2	2	-332	0	-2
711	2694276	2694276	12	12	0	0	-4	-4	384	12	0
731	3357664	3357664	0	0	28	28	4	4	-440	0	4
751	4172746	4172746	-14	-14	-26	-26	-4	-4	470	-14	2

Table 29: McKay–Thompson series $H^{(5)}_{g,4}$

$[g]$	1A	2A	2B	2C	3A	6A	5A	10A	4AB	4CD	12AB
Γ_g	1	$1\vert 4$	$2\vert 2$	2	$3\vert 3$	$3\vert 12$	5	$5\vert 4$	$2\vert 8$	4	$6\vert 24$
4	2	-2	2	-2	2	-2	2	-2	0	0	0
24	12	-12	-4	4	0	0	2	-2	0	0	0
44	20	-20	4	-4	-4	4	0	0	0	0	0
64	50	-50	-6	6	2	-2	0	0	0	0	0
84	72	-72	8	-8	0	0	2	-2	0	0	0
104	152	-152	-8	8	-4	4	2	-2	0	0	0
124	220	-220	12	-12	4	-4	0	0	0	0	0
144	378	-378	-14	14	0	0	-2	2	0	0	0
164	560	-560	16	-16	-4	4	0	0	0	0	0
184	892	-892	-20	20	4	-4	2	-2	0	0	0
204	1272	-1272	24	-24	0	0	2	-2	0	0	0
224	1940	-1940	-28	28	-4	4	0	0	0	0	0
244	2720	-2720	32	-32	8	-8	0	0	0	0	0
264	3960	-3960	-40	40	0	0	0	0	0	0	0
284	5500	-5500	44	-44	-8	8	0	0	0	0	0
304	7772	-7772	-52	52	8	-8	2	-2	0	0	0
324	10590	-10590	62	-62	0	0	0	0	0	0	0
344	14668	-14668	-68	68	-8	8	-2	2	0	0	0
364	19728	-19728	80	-80	12	-12	-2	2	0	0	0
384	26772	-26772	-92	92	0	0	2	-2	0	0	0
404	35624	-35624	104	-104	-16	16	4	-4	0	0	0
424	47592	-47592	-120	120	12	-12	2	-2	0	0	0
444	62568	-62568	136	-136	0	0	-2	2	0	0	0
464	82568	-82568	-152	152	-16	16	-2	2	0	0	0
484	107502	-107502	174	-174	18	-18	2	-2	0	0	0
504	140172	-140172	-196	196	0	0	2	-2	0	0	0
524	180940	-180940	220	-220	-20	20	0	0	0	0	0
544	233576	-233576	-248	248	20	-20	-4	4	0	0	0
564	298968	-298968	280	-280	0	0	-2	2	0	0	0
584	382632	-382632	-312	312	-24	24	2	-2	0	0	0
604	486124	-486124	348	-348	28	-28	4	-4	0	0	0
624	617112	-617112	-392	392	0	0	2	-2	0	0	0
644	778768	-778768	432	-432	-32	32	-2	2	0	0	0
664	981548	-981548	-484	484	32	-32	-2	2	0	0	0
684	1230732	-1230732	540	-540	0	0	2	-2	0	0	0
704	1541244	-1541244	-596	596	-36	36	4	-4	0	0	0
724	1921240	-1921240	664	-664	40	-40	0	0	0	0	0
744	2391456	-2391456	-736	736	0	0	-4	4	0	0	0

C.5 Lambency Seven

Table 30: $H^{(7)}_{g,1}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
-1	-2	-2	-2	-2	-2
27	4	4	4	1	1
55	6	6	-2	0	0
83	10	10	2	-2	-2
111	20	20	-4	2	2
139	30	30	6	0	0
167	42	42	-6	0	0
195	68	68	4	2	2
223	96	96	-8	0	0
251	130	130	10	-2	-2
279	188	188	-12	2	2
307	258	258	10	0	0
335	350	350	-10	-4	-4
363	474	474	18	3	3
391	624	624	-16	0	0
419	826	826	18	-2	-2
447	1090	1090	-22	4	4
475	1410	1410	26	0	0
503	1814	1814	-26	-4	-4
531	2338	2338	26	4	4
559	2982	2982	-34	0	0
587	3774	3774	38	-6	-6
615	4774	4774	-42	4	4
643	5994	5994	42	0	0
671	7494	7494	-50	-6	-6
699	9348	9348	60	6	6
727	11586	11586	-62	0	0
755	14320	14320	64	-8	-8
783	17654	17654	-74	8	8
811	21654	21654	86	0	0
839	26488	26488	-88	-8	-8
867	32334	32334	94	9	9
895	39324	39324	-108	0	0
923	47680	47680	120	-8	-8
951	57688	57688	-128	10	10
979	69600	69600	136	0	0
1007	83760	83760	-152	-12	-12
1035	100596	100596	172	12	12

Table 31: $H^{(7)}_{g,2}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
24	4	-4	0	-2	2
52	12	-12	0	0	0
80	20	-20	0	2	-2
108	32	-32	0	-1	1
136	48	-48	0	0	0
164	80	-80	0	2	-2
192	108	-108	0	-3	3
220	168	-168	0	0	0
248	232	-232	0	4	-4
276	328	-328	0	-2	2
304	444	-444	0	0	0
332	620	-620	0	2	-2
360	812	-812	0	-4	4
388	1104	-1104	0	0	0
416	1444	-1444	0	4	-4
444	1904	-1904	0	-4	4
472	2460	-2460	0	0	0
500	3208	-3208	0	4	-4
528	4080	-4080	0	-6	6
556	5244	-5244	0	0	0
584	6632	-6632	0	8	-8
612	8400	-8400	0	-6	6
640	10524	-10524	0	0	0
668	13224	-13224	0	6	-6
696	16408	-16408	0	-8	8
724	20436	-20436	0	0	0
752	25216	-25216	0	10	-10
780	31120	-31120	0	-8	8
808	38148	-38148	0	0	0
836	46784	-46784	0	8	-8
864	56976	-56976	0	-12	12
892	69432	-69432	0	0	0
920	84144	-84144	0	12	-12
948	101904	-101904	0	-12	12
976	122868	-122868	0	0	0
1004	148076	-148076	0	14	-14
1032	177656	-177656	0	-16	16
1060	213072	-213072	0	0	0

Table 32: $H^{(7)}_{g,3}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
19	6	6	-2	0	0
47	12	12	4	0	0
75	22	22	-2	1	1
103	36	36	4	0	0
131	58	58	-6	-2	-2
159	90	90	2	0	0
187	132	132	-4	0	0
215	190	190	6	-2	-2
243	274	274	-6	1	1
271	384	384	8	0	0
299	528	528	-8	0	0
327	722	722	10	2	2
355	972	972	-12	0	0
383	1300	1300	12	-2	-2
411	1724	1724	-12	2	2
439	2256	2256	16	0	0
467	2938	2938	-22	-2	-2
495	3806	3806	22	2	2
523	4890	4890	-22	0	0
551	6244	6244	28	-2	-2
579	7940	7940	-28	2	2
607	10038	10038	30	0	0
635	12620	12620	-36	-4	-4
663	15814	15814	38	4	4
691	19722	19722	-46	0	0
719	24490	24490	50	-2	-2
747	30310	30310	-50	4	4
775	37362	37362	58	0	0
803	45908	45908	-68	-4	-4
831	56236	56236	68	4	4
859	68646	68646	-74	0	0
887	83556	83556	84	-6	-6
915	101436	101436	-92	6	6
943	122790	122790	102	0	0
971	148254	148254	-106	-6	-6
999	178566	178566	118	6	6
1027	214548	214548	-132	0	0
1055	257190	257190	142	-6	-6

Table 33: $H^{(7)}_{g,4}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
12	4	-4	0	1	-1
40	12	-12	0	0	0
68	16	-16	0	-2	2
96	36	-36	0	0	0
124	48	-48	0	0	0
152	84	-84	0	0	0
180	116	-116	0	2	-2
208	180	-180	0	0	0
236	244	-244	0	-2	2
264	360	-360	0	0	0
292	480	-480	0	0	0
320	676	-676	0	-2	2
348	896	-896	0	2	-2
376	1224	-1224	0	0	0
404	1588	-1588	0	-2	2
432	2128	-2128	0	1	-1
460	2736	-2736	0	0	0
488	3588	-3588	0	0	0
516	4576	-4576	0	4	-4
544	5904	-5904	0	0	0
572	7448	-7448	0	-4	4
600	9500	-9500	0	2	-2
628	11892	-11892	0	0	0
656	14992	-14992	0	-2	2
684	18628	-18628	0	4	-4
712	23256	-23256	0	0	0
740	28688	-28688	0	-4	4
768	35532	-35532	0	3	-3
796	43560	-43560	0	0	0
824	53528	-53528	0	-4	4
852	65256	-65256	0	6	-6
880	79656	-79656	0	0	0
908	96564	-96564	0	-6	6
936	117196	-117196	0	4	-4
964	141360	-141360	0	0	0
992	170600	-170600	0	-4	4
1020	204848	-204848	0	8	-8
1048	245988	-245988	0	0	0

Table 34: $H^{(7)}_{g,5}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
3	2	2	2	-1	-1
31	6	6	-2	0	0
59	14	14	-2	2	2
87	22	22	-2	-2	-2
115	36	36	4	0	0
143	56	56	0	2	2
171	82	82	2	-2	-2
199	126	126	-2	0	0
227	182	182	6	2	2
255	250	250	-6	-2	-2
283	354	354	2	0	0
311	490	490	-6	4	4
339	656	656	8	-4	-4
367	882	882	-6	0	0
395	1180	1180	4	4	4
423	1550	1550	-10	-4	-4
451	2028	2028	12	0	0
479	2638	2638	-10	4	4
507	3394	3394	10	-5	-5
535	4362	4362	-14	0	0
563	5562	5562	18	6	6
591	7032	7032	-16	-6	-6
619	8886	8886	14	0	0
647	11166	11166	-18	6	6
675	13940	13940	28	-7	-7
703	17358	17358	-26	0	0
731	21536	21536	24	8	8
759	26594	26594	-30	-10	-10
787	32742	32742	38	0	0
815	40180	40180	-36	10	10
843	49124	49124	36	-10	-10
871	59916	59916	-44	0	0
899	72852	72852	52	12	12
927	88296	88296	-56	-12	-12
955	106788	106788	52	0	0
983	128816	128816	-64	14	14
1011	154948	154948	76	-14	-14

Table 35: $H^{(7)}_{g,6}$

$[g]$	1A	2A	4A	3AB	6AB
Γ_g	1	1|4	2|8	3	3|4
20	4	-4	0	-2	2
48	4	-4	0	1	-1
76	12	-12	0	0	0
104	12	-12	0	0	0
132	32	-32	0	2	-2
160	36	-36	0	0	0
188	64	-64	0	-2	2
216	80	-80	0	2	-2
244	132	-132	0	0	0
272	160	-160	0	-2	2
300	252	-252	0	3	-3
328	312	-312	0	0	0
356	448	-448	0	-2	2
384	572	-572	0	2	-2
412	792	-792	0	0	0
440	992	-992	0	-4	4
468	1348	-1348	0	4	-4
496	1680	-1680	0	0	0
524	2220	-2220	0	-6	6
552	2776	-2776	0	4	-4
580	3600	-3600	0	0	0
608	4460	-4460	0	-4	4
636	5712	-5712	0	6	-6
664	7044	-7044	0	0	0
692	8892	-8892	0	-6	6
720	10932	-10932	0	6	-6
748	13656	-13656	0	0	0
776	16672	-16672	0	-8	8
804	20672	-20672	0	8	-8
832	25116	-25116	0	0	0
860	30856	-30856	0	-8	8
888	37352	-37352	0	8	-8
916	45564	-45564	0	0	0
944	54884	-54884	0	-10	10
972	66572	-66572	0	11	-11
1000	79848	-79848	0	0	0
1028	96256	-96256	0	-14	14

C.6 Lambency Thirteen

Table 36: $H^{(13)}_{g,1}$

$[g]$	1A	2A	4AB
Γ_g	1	$1 \vert 4$	$2 \vert 8$
-1	-2	-2	-2
51	2	2	2
103	2	2	-2
155	0	0	0
207	2	2	-2
259	2	2	2
311	4	4	0
363	6	6	2
415	6	6	-2
467	8	8	4
519	12	12	-4
571	14	14	2
623	14	14	-2
675	20	20	4
727	24	24	-4
779	28	28	4
831	36	36	-4
883	42	42	6
935	50	50	-6
987	62	62	6
1039	70	70	-6
1091	84	84	8
1143	102	102	-6
1195	118	118	6
1247	136	136	-8
1299	162	162	10
1351	190	190	-10
1403	216	216	8
1455	254	254	-10
1507	292	292	12
1559	336	336	-12
1611	392	392	12
1663	446	446	-14
1715	510	510	14
1767	592	592	-16
1819	672	672	16
1871	764	764	-16
1923	876	876	20

Table 37: $H^{(13)}_{g,2}$

$[g]$	1A	2A	4AB
Γ_g	1	$1 \vert 4$	$2 \vert 8$
48	0	0	0
100	2	-2	0
152	4	-4	0
204	4	-4	0
256	6	-6	0
308	8	-8	0
360	8	-8	0
412	12	-12	0
464	16	-16	0
516	20	-20	0
568	24	-24	0
620	32	-32	0
672	36	-36	0
724	48	-48	0
776	56	-56	0
828	68	-68	0
880	80	-80	0
932	100	-100	0
984	112	-112	0
1036	140	-140	0
1088	164	-164	0
1140	192	-192	0
1192	224	-224	0
1244	268	-268	0
1296	306	-306	0
1348	364	-364	0
1400	420	-420	0
1452	488	-488	0
1504	560	-560	0
1556	656	-656	0
1608	744	-744	0
1660	864	-864	0
1712	988	-988	0
1764	1134	-1134	0
1816	1292	-1292	0
1868	1484	-1484	0
1920	1676	-1676	0
1972	1920	-1920	0

Table 38: $H^{(13)}_{g,3}$

$[g]$	1A	2A	4AB
Γ_g	1	$1 \vert 4$	$2 \vert 8$
43	2	2	-2
95	2	2	2
147	4	4	0
199	6	6	2
251	8	8	-4
303	10	10	2
355	14	14	-2
407	18	18	2
459	22	22	-2
511	26	26	2
563	34	34	-2
615	44	44	4
667	52	52	-4
719	64	64	4
771	78	78	-2
823	96	96	4
875	114	114	-6
927	136	136	4
979	164	164	-4
1031	194	194	6
1083	230	230	-6
1135	270	270	6
1187	318	318	-6
1239	374	374	6
1291	434	434	-10
1343	506	506	10
1395	592	592	-8
1447	686	686	10
1499	792	792	-12
1551	914	914	10
1603	1054	1054	-10
1655	1214	1214	14
1707	1394	1394	-14
1759	1594	1594	14
1811	1822	1822	-14
1863	2084	2084	16
1915	2374	2374	-18
1967	2698	2698	18

Table 39: $H^{(13)}_{g,4}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
36	2	-2	0
88	4	-4	0
140	4	-4	0
192	8	-8	0
244	8	-8	0
296	12	-12	0
348	16	-16	0
400	22	-22	0
452	24	-24	0
504	36	-36	0
556	40	-40	0
608	52	-52	0
660	64	-64	0
712	80	-80	0
764	92	-92	0
816	116	-116	0
868	136	-136	0
920	168	-168	0
972	196	-196	0
1024	238	-238	0
1076	272	-272	0
1128	332	-332	0
1180	384	-384	0
1232	456	-456	0
1284	528	-528	0
1336	620	-620	0
1388	712	-712	0
1440	840	-840	0
1492	960	-960	0
1544	1120	-1120	0
1596	1280	-1280	0
1648	1484	-1484	0
1700	1688	-1688	0
1752	1952	-1952	0
1804	2216	-2216	0
1856	2544	-2544	0
1908	2888	-2888	0
1960	3304	-3304	0

Table 40: $H^{(13)}_{g,5}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
27	2	2	2
79	4	4	0
131	6	6	2
183	6	6	-2
235	10	10	2
287	14	14	-2
339	16	16	0
391	22	22	-2
443	30	30	2
495	36	36	-4
547	46	46	2
599	58	58	-2
651	68	68	4
703	86	86	-2
755	106	106	2
807	124	124	-4
859	152	152	4
911	184	184	-4
963	216	216	4
1015	258	258	-6
1067	308	308	4
1119	362	362	-6
1171	426	426	6
1223	502	502	-6
1275	584	584	8
1327	684	684	-8
1379	798	798	6
1431	920	920	-8
1483	1070	1070	10
1535	1238	1238	-10
1587	1422	1422	10
1639	1638	1638	-10
1691	1884	1884	12
1743	2156	2156	-12
1795	2468	2468	12
1847	2822	2822	-14
1899	3212	3212	16
1951	3660	3660	-16

Table 41: $H^{(13)}_{g,6}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
16	2	-2	0
68	4	-4	0
120	4	-4	0
172	8	-8	0
224	8	-8	0
276	16	-16	0
328	16	-16	0
380	24	-24	0
432	28	-28	0
484	38	-38	0
536	44	-44	0
588	60	-60	0
640	68	-68	0
692	88	-88	0
744	104	-104	0
796	132	-132	0
848	152	-152	0
900	190	-190	0
952	220	-220	0
1004	268	-268	0
1056	312	-312	0
1108	376	-376	0
1160	432	-432	0
1212	520	-520	0
1264	596	-596	0
1316	708	-708	0
1368	812	-812	0
1420	956	-956	0
1472	1092	-1092	0
1524	1280	-1280	0
1576	1460	-1460	0
1628	1696	-1696	0
1680	1932	-1932	0
1732	2236	-2236	0
1784	2536	-2536	0
1836	2924	-2924	0
1888	3308	-3308	0
1940	3792	-3792	0

Table 42: $H^{(13)}_{g,7}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
3	2	2	-2
55	2	2	2
107	4	4	0
159	8	8	0
211	10	10	-2
263	12	12	0
315	16	16	0
367	22	22	2
419	26	26	-2
471	34	34	2
523	44	44	0
575	54	54	2
627	68	68	-4
679	82	82	2
731	102	102	-2
783	124	124	4
835	148	148	-4
887	176	176	4
939	214	214	-2
991	256	256	4
1043	300	300	-4
1095	356	356	4
1147	420	420	-4
1199	494	494	6
1251	580	580	-8
1303	674	674	6
1355	786	786	-6
1407	918	918	6
1459	1060	1060	-8
1511	1226	1226	6
1563	1418	1418	-6
1615	1632	1632	8
1667	1874	1874	-10
1719	2150	2150	10
1771	2464	2464	-8
1823	2816	2816	12
1875	3214	3214	-14

Table 43: $H^{(13)}_{g,8}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
40	4	-4	0
92	4	-4	0
144	6	-6	0
196	6	-6	0
248	12	-12	0
300	12	-12	0
352	20	-20	0
404	24	-24	0
456	32	-32	0
508	36	-36	0
560	52	-52	0
612	56	-56	0
664	76	-76	0
716	88	-88	0
768	112	-112	0
820	128	-128	0
872	164	-164	0
924	184	-184	0
976	232	-232	0
1028	268	-268	0
1080	324	-324	0
1132	372	-372	0
1184	452	-452	0
1236	512	-512	0
1288	616	-616	0
1340	704	-704	0
1392	832	-832	0
1444	950	-950	0
1496	1120	-1120	0
1548	1268	-1268	0
1600	1486	-1486	0
1652	1688	-1688	0
1704	1956	-1956	0
1756	2220	-2220	0
1808	2568	-2568	0
1860	2896	-2896	0
1912	3336	-3336	0

Table 44: $H^{(13)}_{g,9}$

$[g]$	1A	2A	4AB
Γ_g	1	1|4	2|8
23	2	2	-2
75	4	4	0
127	4	4	0
179	6	6	2
231	8	8	0
283	12	12	0
335	14	14	-2
387	20	20	0
439	26	26	-2
491	30	30	2
543	40	40	0
595	50	50	2
647	60	60	0
699	74	74	2
751	90	90	-2
803	108	108	4
855	134	134	-2
907	158	158	2
959	188	188	-4
1011	226	226	2
1063	266	266	-2
1115	314	314	2
1167	372	372	-4
1219	436	436	4
1271	508	508	-4
1323	596	596	4
1375	692	692	-4
1427	802	802	6
1479	932	932	-4
1531	1074	1074	6
1583	1238	1238	-6
1635	1430	1430	6
1687	1640	1640	-8
1739	1878	1878	6
1791	2150	2150	-6
1843	2456	2456	8
1895	2800	2800	-8

Table 45: $H^{(13)}_{g,10}$

$[g]$	1A	2A	4AB
Γ_g	1	$1\mid 4$	$2\mid 8$
4	2	-2	0
56	0	0	0
108	4	-4	0
160	4	-4	0
212	8	-8	0
264	8	-8	0
316	12	-12	0
368	12	-12	0
420	20	-20	0
472	20	-20	0
524	32	-32	0
576	34	-34	0
628	48	-48	0
680	52	-52	0
732	72	-72	0
784	78	-78	0
836	104	-104	0
888	116	-116	0
940	148	-148	0
992	164	-164	0
1044	208	-208	0
1096	232	-232	0
1148	288	-288	0
1200	324	-324	0
1252	396	-396	0
1304	444	-444	0
1356	536	-536	0
1408	604	-604	0
1460	720	-720	0
1512	812	-812	0
1564	960	-960	0
1616	1080	-1080	0
1668	1268	-1268	0
1720	1428	-1428	0
1772	1664	-1664	0
1824	1872	-1872	0

Table 46: $H^{(13)}_{g,11}$

$[g]$	1A	2A	4AB
Γ_g	1	$1\mid 4$	$2\mid 8$
35	2	2	2
87	2	2	2
139	2	2	-2
191	4	4	0
243	4	4	0
295	8	8	0
347	10	10	-2
399	10	10	2
451	16	16	0
503	20	20	0
555	22	22	-2
607	28	28	0
659	36	36	0
711	44	44	0
763	54	54	-2
815	64	64	0
867	76	76	0
919	94	94	2
971	114	114	-2
1023	130	130	2
1075	156	156	0
1127	188	188	0
1179	216	216	-4
1231	254	254	2
1283	300	300	0
1335	346	346	2
1387	404	404	-4
1439	470	470	2
1491	542	542	-2
1543	630	630	2
1595	724	724	-4
1647	828	828	4
1699	954	954	-2
1751	1100	1100	4
1803	1250	1250	-6
1855	1428	1428	4

Table 47: $H^{(13)}_{g,12}$

$[g]$	1A	2A	4AB
Γ_g	1	$1\mid 4$	$2\mid 8$
12	0	0	0
64	2	-2	0
116	0	0	0
168	4	-4	0
220	0	0	0
272	4	-4	0
324	2	-2	0
376	8	-8	0
428	4	-4	0
480	12	-12	0
532	8	-8	0
584	16	-16	0
636	16	-16	0
688	24	-24	0
740	20	-20	0
792	36	-36	0
844	32	-32	0
896	48	-48	0
948	48	-48	0
1000	68	-68	0
1052	68	-68	0
1104	96	-96	0
1156	98	-98	0
1208	128	-128	0
1260	136	-136	0
1312	176	-176	0
1364	184	-184	0
1416	240	-240	0
1468	252	-252	0
1520	312	-312	0
1572	340	-340	0
1624	416	-416	0
1676	448	-448	0
1728	548	-548	0
1780	592	-592	0
1832	708	-708	0

D Decompositions

As explained in §4 (see also §5.1) our conjectural proposals for the umbral McKay–Thompson series $H^{(\ell)}_{g,r}(\tau)=\sum_d c^{(\ell)}_{g,r}(d)q^d$ (cf. §§4,C) determine the $G^{(\ell)}$-modules $K^{(\ell)}_{r,d}$ up to isomorphism for $d>0$, at least for those values of d for which we can identify all the Fourier coefficients $c^{(\ell)}_{g,r}(d)$. In this section we furnish tables of explicit decompositions into irreducible representations of $G^{(\ell)}$ for $K^{(\ell)}_{r,d}$, for the first few values of d. The coefficient $c^{(\ell)}_r(d)$ of $H^{(\ell)}_r=H^{(\ell)}_{e,r}$ is non-zero only when $d=n-r^2/4\ell$ for some integer $n\geq 0$. For each of the tables in this section the rows are labelled by the values $4\ell d$, so that the entry in row m and column χ_i indicates the multiplicity of the irreducible representation of $G^{(\ell)}$ with character χ_i (in the notation of the character tables of §B.1) appearing in the $G^{(\ell)}$-module $K^{(\ell)}_{r,m/4\ell}$. One can observe that these tables support Conjectures 5.1, 5.11 and 5.12, and also give evidence in support of the hypothesis that $K^{(\ell)}_{r,d}$ has a decomposition into irreducible representations that factor through $\bar{G}^{(\ell)}$ when r is odd, and has a decomposition into faithful irreducible representations of $G^{(\ell)}$ when r is even.

D.1 Lambency Two

Table 48: Decomposition of $K_1^{(2)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	χ_{15}	χ_{16}	χ_{17}	χ_{18}	χ_{19}	χ_{20}	χ_{21}	χ_{22}	χ_{23}	χ_{24}	χ_{25}	χ_{26}
-1	-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	2	2	2	2
63	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	2	0	0	2	2	2	4	2	2	6
71	0	0	0	0	0	0	0	0	2	2	2	0	0	2	2	2	0	2	2	2	4	4	4	8	8	10
79	0	0	0	0	2	2	0	2	2	0	0	2	2	2	2	2	4	4	4	6	6	8	12	10	10	24
87	0	0	0	0	0	0	0	0	0	4	4	4	4	6	4	4	2	8	10	8	14	12	22	24	26	40
95	0	2	0	0	2	2	2	4	4	6	6	8	8	4	8	8	12	12	12	18	26	30	40	38	40	80
103	0	0	2	2	2	2	4	2	6	10	10	14	14	18	14	14	16	26	30	28	44	44	70	80	84	136
111	0	0	0	0	8	8	4	6	14	16	16	24	24	22	24	24	34	38	46	58	80	86	128	126	132	254
119	0	0	2	2	8	8	12	8	18	38	38	40	40	46	44	44	46	78	86	88	138	144	218	238	246	424
127	0	2	2	2	18	18	18	22	36	50	50	72	72	68	72	72	100	122	140	170	232	252	378	382	400	742
135	0	2	8	8	25	25	30	26	54	94	94	116	116	130	124	124	140	212	246	262	392	410	630	670	704	1222
143	0	6	6	6	50	50	50	58	100	148	148	194	194	192	202	202	256	342	388	454	654	704	1044	1074	1120	2058
151	0	4	18	18	68	68	80	72	150	252	252	318	318	346	332	332	394	582	664	722	1062	1116	1702	1800	1880	3320
159	0	14	20	20	126	126	128	138	254	390	390	516	516	520	536	536	676	904	1036	1196	1716	1836	2764	2846	2980	5408
167	2	20	40	40	182	182	214	200	396	652	652	814	814	872	860	860	1020	1476	1684	1862	2742	2902	4384	4622	4828	8572
175	2	32	55	55	314	314	328	346	640	988	988	1298	1298	1336	1348	1348	1686	2302	2630	3000	4324	4616	6950	7204	7532	13620
183	2	40	98	98	460	460	512	496	972	1590	1590	2020	2020	2144	2118	2118	2546	3638	4162	4624	6768	7166	10856	11376	11898	21204

D.2 Lambency Three

Table 49: Decomposition of $K_1^{(3)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	χ_{15}
-1	-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
47	0	0	0	0	0	0	0	0	0	2	0	2	0	0	2
59	0	0	0	0	0	2	2	0	0	0	2	2	2	2	2
71	0	0	2	0	0	0	0	2	2	2	2	2	4	4	6
83	0	0	0	2	2	4	4	2	2	2	4	6	6	8	10
95	0	0	2	0	0	4	4	8	8	6	6	10	12	14	18
107	0	2	2	4	4	8	12	8	8	8	14	16	22	28	30

Table 50: Decomposition of $K_2^{(3)}$

	χ_{16}	χ_{17}	χ_{18}	χ_{19}	χ_{20}	χ_{21}	χ_{22}	χ_{23}	χ_{24}	χ_{25}	χ_{26}
8	1	1	0	0	0	0	0	0	0	0	0
20	0	0	0	0	1	1	0	0	0	0	0
32	0	0	0	0	0	0	1	1	0	0	0
44	0	0	0	0	0	0	0	0	2	1	1
56	0	0	0	2	0	0	2	2	0	2	2
68	0	0	2	0	2	2	2	2	4	4	4
80	2	2	0	0	1	1	6	6	4	8	8
92	0	0	2	4	6	6	8	8	12	14	14
104	2	2	0	4	6	6	20	20	16	24	24
116	2	2	6	8	12	12	26	26	36	44	44

D.3 Lambency Four

Table 51: Decomposition of $K_1^{(4)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}
-1	-2	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	2	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	2
47	0	0	0	0	0	2	0	0	2	2	0
63	0	1	1	0	2	0	0	2	2	2	4
79	0	0	0	2	2	2	4	0	4	6	4
95	0	2	2	2	2	4	2	4	6	8	12
111	2	2	2	6	6	6	8	6	10	18	14
127	0	4	4	6	10	10	6	12	18	26	30
143	2	6	6	14	14	18	18	10	32	46	40
159	4	10	10	18	24	26	20	26	44	68	76

Table 52: Decomposition of $K_3^{(4)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}
7	0	1	1	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	2	0	0
39	0	0	0	0	2	0	0	0	0	2	0
55	0	0	0	2	0	0	0	2	2	2	2
71	0	2	2	0	2	2	0	0	2	4	4
87	2	0	0	2	2	2	4	4	6	6	8
103	0	2	2	2	6	6	4	2	6	14	12
119	2	2	2	8	4	8	6	8	18	20	22
135	2	8	8	8	14	14	12	10	20	36	34
151	4	6	6	18	16	20	20	22	42	54	56

Table 53: Decomposition of $K_2^{(4)}$

	χ_{12}	χ_{13}	χ_{14}	χ_{15}	χ_{16}
12	0	1	1	0	0
28	0	0	0	1	1
44	2	0	0	2	2
60	0	2	2	4	4
76	2	2	2	8	8
92	6	4	4	14	14
108	6	9	9	24	24
124	14	14	14	40	40
140	24	20	20	66	66
156	32	36	36	104	104

D.4 Lambency Five

Table 54: Decomposition of $K_1^{(5)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7
-1	-2	0	0	0	0	0	0
19	0	0	2	0	0	0	0
39	0	0	0	2	0	2	0
59	0	0	2	0	2	2	2
79	0	2	2	4	2	2	4
99	2	0	6	2	6	4	6
119	0	2	6	8	8	10	10
139	4	2	14	10	14	12	16
159	2	6	14	20	20	22	26
179	8	4	28	22	36	32	40

Table 55: Decomposition of $K_3^{(5)}$

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7
11	0	0	0	2	0	0	0
31	0	0	0	0	2	0	2
51	0	0	2	2	2	2	2
71	2	0	4	2	4	4	4
91	0	2	4	6	6	8	8
111	2	2	10	8	12	10	14
131	4	4	14	16	18	18	22
151	8	4	24	22	30	30	34
171	8	10	34	38	44	46	54
191	14	14	58	52	68	64	82

Table 56: Decomposition of $K_2^{(5)}$

	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}
16	0	0	0	0	1	1	0
36	0	0	1	1	1	1	2
56	2	2	2	2	2	2	2
76	0	0	4	4	4	4	6
96	2	2	6	6	8	8	8
116	2	2	10	10	12	12	18
136	4	4	16	16	22	22	22
156	6	6	26	26	32	32	42
176	12	12	40	40	50	50	56
196	13	13	60	60	74	74	94

Table 57: Decomposition of $K_4^{(5)}$

	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}
4	1	1	0	0	0	0	0
24	0	0	0	0	0	0	2
44	0	0	0	0	2	2	0
64	0	0	2	2	1	1	4
84	2	2	2	2	4	4	2
104	0	0	4	4	6	6	10
124	4	4	8	8	10	10	8
144	1	1	13	13	14	14	22
164	6	6	18	18	26	26	24
184	6	6	30	30	34	34	50

D.5 Lambency Seven

Table 58: Decomposition of $K_1^{(7)}$

	χ_1	χ_2	χ_3	χ_4
-1	-2	0	0	0
27	2	1	1	0
55	0	0	0	2
83	0	2	2	2
111	2	0	0	6
139	4	4	4	6
167	2	2	2	12
195	8	6	6	16
223	6	6	6	26
251	12	14	14	30

Table 59: Decomposition of $K_3^{(7)}$

	χ_1	χ_2	χ_3	χ_4
19	0	0	0	2
47	2	2	2	2
75	2	1	1	6
103	4	4	4	8
131	2	4	4	16
159	8	8	8	22
187	10	10	10	34
215	16	18	18	46
243	22	21	21	70
271	34	34	34	94

Table 60: Decomposition of $K_5^{(7)}$

	χ_1	χ_2	χ_3	χ_4
3	0	1	1	0
31	0	0	0	2
59	2	0	0	4
87	0	2	2	6
115	4	4	4	8
143	6	4	4	14
171	6	8	8	20
199	10	10	10	32
227	18	16	16	44
255	18	20	20	64

Table 61: Decomposition of $K_2^{(7)}$

	χ_5	χ_6	χ_7
24	2	0	0
52	2	2	2
80	2	4	4
108	6	5	5
136	8	8	8
164	12	14	14
192	20	17	17
220	28	28	28
248	36	40	40
276	56	54	54

Table 62: Decomposition of $K_4^{(7)}$

	χ_5	χ_6	χ_7
12	0	1	1
40	2	2	2
68	4	2	2
96	6	6	6
124	8	8	8
152	14	14	14
180	18	20	20
208	30	30	30
236	42	40	40
264	60	60	60

Table 63: Decomposition of $K_6^{(7)}$

	χ_5	χ_6	χ_7
20	2	0	0
48	0	1	1
76	2	2	2
104	2	2	2
132	4	6	6
160	6	6	6
188	12	10	10
216	12	14	14
244	22	22	22
272	28	26	26

D.6 Lambency Thirteen

Table 64: Decomposition of $K_1^{(13)}$

	χ_1	χ_2
-1	-2	0
51	2	0
103	0	2
155	0	0
207	0	2
259	2	0
311	2	2
363	4	2
415	2	4
467	6	2

Table 67: Decomposition of $K_2^{(13)}$

	χ_3	χ_4
48	0	0
100	1	1
152	2	2
204	2	2
256	3	3
308	4	4
360	4	4
412	6	6
464	8	8
516	10	10

Table 65: Decomposition of $K_3^{(13)}$

	χ_1	χ_2
43	0	2
95	2	0
147	2	2
199	4	2
251	2	6
303	6	4
355	6	8
407	10	8
459	10	12
511	14	12

Table 68: Decomposition of $K_4^{(13)}$

	χ_3	χ_4
36	1	1
88	2	2
140	2	2
192	4	4
244	4	4
296	6	6
348	8	8
400	11	11
452	12	12
504	18	18

Table 66: Decomposition of $K_5^{(13)}$

	χ_1	χ_2
27	2	0
79	2	2
131	4	2
183	2	4
235	6	4
287	6	8
339	8	8
391	10	12
443	16	14
495	16	20

Table 69: Decomposition of $K_6^{(13)}$

	χ_3	χ_4
16	1	1
68	2	2
120	2	2
172	4	4
224	4	4
276	8	8
328	8	8
380	12	12
432	14	14
484	19	19

Table 70: Decomposition of $K_7^{(13)}$

	χ_1	χ_2
3	0	2
55	2	0
107	2	2
159	4	4
211	4	6
263	6	6
315	8	8
367	12	10
419	12	14
471	18	16

Table 71: Decomposition of $K_9^{(13)}$

	χ_1	χ_2
23	0	2
75	2	2
127	2	2
179	4	2
231	4	4
283	6	6
335	6	8
387	10	10
439	12	14
491	16	14

Table 72: Decomposition of $K_{11}^{(13)}$

	χ_1	χ_2
35	2	0
87	2	0
139	0	2
191	2	2
243	2	2
295	4	4
347	4	6
399	6	4
451	8	8
503	10	10

Table 73: Decomposition of $K_8^{(13)}$

	χ_3	χ_4
40	2	2
92	2	2
144	3	3
196	3	3
248	6	6
300	6	6
352	10	10
404	12	12
456	16	16
508	18	18

Table 74: Decomposition of $K_{10}^{(13)}$

	χ_3	χ_4
4	1	1
56	0	0
108	2	2
160	2	2
212	4	4
264	4	4
316	6	6
368	6	6
420	10	10
472	10	10

Table 75: Decomposition of $K_{12}^{(13)}$

	χ_3	χ_4
12	0	0
64	1	1
116	0	0
168	2	2
220	0	0
272	2	2
324	1	1
376	4	4
428	2	2
480	6	6

References

[1] J. H. Conway and S. P. Norton, “Monstrous Moonshine,” *Bull. London Math. Soc.* **11** (1979) 308 339.

[2] J. G. Thompson, “Some numerology between the Fischer-Griess Monster and the elliptic modular function,” *Bull. London Math. Soc.* **11** (1979) no. 3, 352–353.

[3] J. G. Thompson, “Finite groups and modular functions,” *Bull. London Math. Soc.* **11** (1979) no. 3, 347–351.

[4] P. Fong, “Characters arising in the Monster-modular connection,” in *The Santa Cruz Conference on Finite Groups (Univ. California, Santa Cruz, Calif., 1979)*, vol. 37 of *Proc. Sympos. Pure Math.*, pp. 557–559. Amer. Math. Soc., Providence, R.I., 1980.

[5] S. D. Smith, “On the head characters of the Monster simple group,” in *Finite groups—coming of age (Montreal, Que., 1982)*, vol. 45 of *Contemp. Math.*, pp. 303–313. Amer. Math. Soc., Providence, RI, 1985. `http://dx.doi.org/10.1090/conm/045/822245`.

[6] I. B. Frenkel, J. Lepowsky, and A. Meurman, “A natural representation of the Fischer-Griess Monster with the modular function J as character,” *Proc. Nat. Acad. Sci. U.S.A.* **81** (1984) no. 10, Phys. Sci., 3256–3260.

[7] I. B. Frenkel, J. Lepowsky, and A. Meurman, “A moonshine module for the Monster,” in *Vertex operators in mathematics and physics (Berkeley, Calif., 1983)*, vol. 3 of *Math. Sci. Res. Inst. Publ.*, pp. 231–273. Springer, New York, 1985.

[8] I. B. Frenkel and V. G. Kac, “Basic representations of affine Lie algebras and dual resonance models,” *Invent. Math.* **62** (1980/81) no. 1, 23–66.

[9] G. Segal, “Unitary representations of some infinite-dimensional groups,” *Comm. Math. Phys.* **80** (1981) no. 3, 301–342. `http://projecteuclid.org/getRecord?id=euclid.cmp/1103919978`.

[10] R. L. Griess, Jr., “The friendly giant,” *Invent. Math.* **69** (1982) no. 1, 1–102.

[11] R. Borcherds, “Vertex algebras, Kac-Moody algebras, and the Monster,” *Proceedings of the National Academy of Sciences, U.S.A.* **83** (1986) no. 10, 3068–3071.

[12] I. Frenkel, J. Lepowsky, and A. Meurman, *Vertex operator algebras and the Monster*, vol. 134 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1988.

[13] R. Borcherds, “Generalized Kac-Moody algebras,” *J. Algebra* **115** (1988) no. 2, 501–512. `http://dx.doi.org/10.1016/0021-8693(88)90275-X`.

[14] R. E. Borcherds, "Monstrous moonshine and monstrous Lie superalgebras," *Invent. Math.* **109, No.2** (1992) 405–444.

[15] T. Eguchi, H. Ooguri, and Y. Tachikawa, "Notes on the K3 Surface and the Mathieu group M_{24}," *Exper.Math.* **20** (2011) 91–96, `arXiv:1004.0956 [hep-th]`.

[16] T. Eguchi and A. Taormina, "Unitary representations of the $N = 4$ superconformal algebra," *Phys. Lett. B* **196** (1987) no. 1, 75–81. `http://dx.doi.org/10.1016/0370-2693(87)91679-0`.

[17] T. Eguchi and A. Taormina, "Character formulas for the $N = 4$ superconformal algebra," *Phys. Lett. B* **200** (1988) no. 3, 315–322. `http://dx.doi.org/10.1016/0370-2693(88)90778-2`.

[18] V. G. Kac and M. Wakimoto, "Quantum reduction and representation theory of superconformal algebras," *Adv. Math.* **185** (2004) no. 2, 400–458. `http://dx.doi.org/10.1016/j.aim.2003.12.005`.

[19] M. C. N. Cheng, "$K3$ Surfaces, $N = 4$ Dyons, and the Mathieu Group M_{24}," `1005.5415`.

[20] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu twining characters for K3," *JHEP* **1009** (2010) 058, `arXiv:1006.0221 [hep-th]`. 19 pages.

[21] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu Moonshine in the elliptic genus of K3," *JHEP* **1010** (2010) 062, `arXiv:1008.3778 [hep-th]`.

[22] T. Eguchi and K. Hikami, "Note on Twisted Elliptic Genus of K3 Surface," *Phys.Lett.* **B694** (2011) 446–455, `arXiv:1008.4924 [hep-th]`.

[23] T. Gannon, "Much ado about Mathieu," `arXiv:1211.5531 [math.RT]`.

[24] M. C. N. Cheng and J. F. R. Duncan, "On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine," *Commun. Number Theory Phys.* **6** (2012) no. 3, , `1110.3859`.

[25] J. F. R. Duncan and I. B. Frenkel, "Rademacher sums, moonshine and gravity," *Commun. Number Theory Phys.* **5** (2011) no. 4, 1–128.

[26] H. Rademacher, "The Fourier Series and the Functional Equation of the Absolute Modular Invariant J(τ)," *Amer. J. Math.* **61** (1939) no. 1, 237–248.

[27] E. Witten, "Three-Dimensional Gravity Revisited," `0706.3359`.

[28] A. Maloney and E. Witten, "Quantum gravity partition functions in three dimensions," *J. High Energy Phys.* (2010) no. 2, 029, 58. `http://dx.doi.org/10.1007/JHEP02(2010)029`.

[29] W. Li, W. Song, and A. Strominger, "Chiral gravity in three dimensions," *J. High Energy Phys.* (2008) no. 4, 082, 15. `http://dx.doi.org/10.1088/1126-6708/2008/04/082`.

[30] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," *Adv.Theor.Math.Phys.* **2** (1998) 231–252, `hep-th/9711200`.

[31] S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys.Lett.* **B428** (1998) 105–114, `arXiv:hep-th/9802109 [hep-th]`.

[32] E. Witten, "Anti-de Sitter space and holography," *Adv.Theor.Math.Phys.* **2** (1998) 253–291, `arXiv:hep-th/9802150 [hep-th]`.

[33] R. Dijkgraaf, J. Maldacena, G. Moore, and E. Verlinde, "A black hole farey tail," `hep-th/0005003`.

[34] G. W. Moore, "Les Houches Lectures on Strings and Arithmetic," `hep-th/0401049`.

[35] J. de Boer, M. C. N. Cheng, R. Dijkgraaf, J. Manschot, and E. Verlinde, "A farey tail for attractor black holes," *JHEP* **0611:024,2006** (JHEP0611:024,2006) , `hep-th/0608059`.

[36] P. Kraus and F. Larsen, "Partition functions and elliptic genera from supergravity," *JHEP* **0701:002,2007** (JHEP 0701:002,2007) , `hep-th/0607138`.

[37] F. Denef and G. W. Moore, "Split states, entropy enigmas, holes and halos," *JHEP* **1111** (2011) 129, `arXiv:hep-th/0702146 [HEP-TH]`. 149 pages, 21 figures.

[38] A. Sen, "Black hole entropy function, attractors and precision counting of microstates," *Gen.Rel.Grav.* **40:2249-2431,2008** (Aug., 2007) , `0708.1270`.

[39] A. Dabholkar, J. Gomes, and S. Murthy, "Localization and Exact Holography," `arXiv:1111.1161 [hep-th]`.

[40] J. Manschot and G. W. Moore, "A modern fareytail," *Commun.Num.Theor.Phys.* **4,** (Dec., 2007) 103–159,2010, `0712.0573`.

[41] M. C. N. Cheng and J. F. R. Duncan, "The largest Mathieu group and (mock) automorphic forms," in *String-Math 2011*, vol. 85 of *Proc. Sympos. Pure Math.*, pp. 53–82. Amer. Math. Soc., Providence, RI, 2012. `http://dx.doi.org/10.1090/pspum/085/1374`.

[42] J. S. Ellenberg, "On the error term in Duke's estimate for the average special value of L-functions," *Canad. Math. Bull.* **48** (2005) no. 4, 535–546. `http://dx.doi.org/10.4153/CMB-2005-049-8`.

[43] V. Gritsenko, "Elliptic genus of Calabi-Yau manifolds and Jacobi and Siegel modular forms," *Algebra i Analiz* **11** (1999) no. 5, 100–125.

[44] S. Zwegers, *Mock Theta Functions*. PhD thesis, Utrecht University, 2002.

[45] A. Dabholkar, S. Murthy, and D. Zagier, "Quantum Black Holes, Wall Crossing, and Mock Modular Forms," `arXiv:1208.4074 [hep-th]`.

[46] K. Bringmann and K. Ono, "The $f(q)$ mock theta function conjecture and partition ranks," *Invent. Math.* **165** (2006) no. 2, 243–266. `http://dx.doi.org/10.1007/s00222-005-0493-5`.

[47] K. Bringmann and K. Ono, "Dyson's ranks and Maass forms," *Ann. of Math. (2)* **171** (2010) no. 1, 419–449. `http://dx.doi.org/10.4007/annals.2010.171.419`.

[48] T. Eguchi and K. Hikami, "Superconformal algebras and mock theta functions,"*J.Phys.A* **42:304010,2009** (Dec., 2008) , `0812.1151`.

[49] T. Eguchi and K. Hikami, "Superconformal Algebras and Mock Theta Functions 2. Rademacher Expansion for K3 Surface,"*Communications in Number Theory and Physics* **3,** (Apr., 2009) 531–554, `0904.0911`.

[50] C. Vafa and E. Witten, "A strong coupling test of S-duality," *Nuclear Phys. B* **431** (1994) no. 1-2, 3–77. `http://dx.doi.org/10.1016/0550-3213(94)90097-3`.

[51] L. Göttsche and D. Zagier, "Jacobi forms and the structure of Donaldson invariants for 4-manifolds with $b_+ = 1$," *Selecta Math. (N.S.)* **4** (1998) no. 1, 69–115. `http://dx.doi.org/10.1007/s000290050025`.

[52] J. Troost, "The non-compact elliptic genus: mock or modular," *JHEP* **1006** (2010) 104, `arXiv:1004.3649 [hep-th]`.

[53] J. Manschot, "The Betti numbers of the moduli space of stable sheaves of rank 3 on $\mathbb{P}^2$," *Lett. Math. Phys.* **98** (2011) no. 1, 65–78. `http://dx.doi.org/10.1007/s11005-011-0490-0`.

[54] R. Lawrence and D. Zagier, "Modular forms and quantum invariants of 3-manifolds," *Asian J. Math.* **3** (1999) no. 1, 93–107. Sir Michael Atiyah: a great mathematician of the twentieth century.

[55] D. Zagier, "Ramanujan's mock theta functions and their applications (after Zwegers and Ono-Bringmann)," *Astérisque* (2009) no. 326, Exp. No. 986, vii–viii, 143–164 (2010). Séminaire Bourbaki. Vol. 2007/2008.

[56] A. Folsom, “What is … a mock modular form?,” *Notices Amer. Math. Soc.* **57** (2010) no. 11, 1441–1443.

[57] K. Ono, “Unearthing the visions of a master: harmonic Maass forms and number theory,” in *Current developments in mathematics, 2008*, pp. 347–454. Int. Press, Somerville, MA, 2009.

[58] S. Ramanujan, *The lost notebook and other unpublished papers*. Springer-Verlag, Berlin, 1988. With an introduction by George E. Andrews.

[59] B. Gordon and R. J. McIntosh, “Some eighth order mock theta functions,” *J. London Math. Soc. (2)* **62** (2000) no. 2, 321–335. `http://dx.doi.org/10.1112/S0024610700008735`.

[60] M. Eichler and D. Zagier, *The theory of Jacobi forms*. Birkhäuser, 1985.

[61] V. Gritsenko, “Complex vector bundles and Jacobi forms,” *Sūrikaisekikenkyūsho Kōkyūroku* (1999) no. 1103, 71–85. Automorphic forms and L-functions (Kyoto, 1999).

[62] N.-P. Skoruppa, *Über den Zusammenhang zwischen Jacobiformen und Modulformen halbganzen Gewichts*. Bonner Mathematische Schriften [Bonn Mathematical Publications], 159. Universität Bonn Mathematisches Institut, Bonn, 1985. Dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1984.

[63] A. W. Knapp, *Elliptic curves*, vol. 40 of *Mathematical Notes*. Princeton University Press, Princeton, NJ, 1992.

[64] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*. Publications of the Mathematical Society of Japan, No. 11. Iwanami Shoten, Publishers, Tokyo, 1971. Kanô Memorial Lectures, No. 1.

[65] B. J. Birch and H. P. F. Swinnerton-Dyer, “Notes on elliptic curves. II,” *J. Reine Angew. Math.* **218** (1965) 79–108.

[66] A. Wiles, “The Birch and Swinnerton-Dyer conjecture,” in *The millennium prize problems*, pp. 31–41. Clay Math. Inst., Cambridge, MA, 2006.

[67] J. H. Bruinier and J. Funke, “On two geometric theta lifts,” *Duke Math. J.* **125** (2004) no. 1, 45–90. `http://dx.doi.org/10.1215/S0012-7094-04-12513-8`.

[68] N.-P. Skoruppa, “Explicit formulas for the Fourier coefficients of Jacobi and elliptic modular forms,” *Invent. Math.* **102** (1990) no. 3, 501–520. `http://dx.doi.org/10.1007/BF01233438`.

[69] N.-P. Skoruppa, “Heegner cycles, modular forms and Jacobi forms,” *Sém. Théor. Nombres Bordeaux (2)* **3** (1991) no. 1, 93–116. `http://jtnb.cedram.org/item?id=JTNB_1991__3_1_93_0`.

[70] W. Duke, “The critical order of vanishing of automorphic L-functions with large level,” *Invent. Math.* **119** (1995) no. 1, 165–174. `http://dx.doi.org/10.1007/BF01245178`.

[71] W. Stein, “The Modular Forms Database,”. `http://wstein.org/Tables`.

[72] G. Hoehn, “Selbstduale vertexoperatorsuperalgebren und das babymonster (self-dual vertex operator super algebras and the baby monster),” *Bonner Mathematische Schriften, Vol.* **286,** (June, 2007) 1–85,Bonn1996, `0706.0236`.

[73] M. R. Gaberdiel, S. Gukov, C. A. Keller, G. W. Moore, and H. Ooguri, “Extremal $N = (2, 2)$ 2D conformal field theories and constraints of modularity,” *Commun. Number Theory Phys.* **2** (2008) no. 4, 743–801.

[74] A. Maloney, W. Song, and A. Strominger, “Chiral Gravity, Log Gravity and Extremal CFT,” *Phys.Rev.* **D81** (2010) 064007, `arXiv:0903.4573 [hep-th]`.

[75] S. Carlip, S. Deser, A. Waldron, and D. Wise, “Topologically Massive AdS Gravity,” *Phys.Lett.* **B666** (2008) 272–276, `arXiv:0807.0486 [hep-th]`.

[76] S. Carlip, S. Deser, A. Waldron, and D. Wise, “Cosmological Topologically Massive Gravitons and Photons,” *Class.Quant.Grav.* **26** (2009) 075008, `arXiv:0803.3998 [hep-th]`.

[77] G. Giribet, M. Kleban, and M. Porrati, “Topologically Massive Gravity at the Chiral Point is Not Chiral,” *JHEP* **0810** (2008) 045, `arXiv:0807.4703 [hep-th]`.

[78] D. Grumiller, R. Jackiw, and N. Johansson, “Canonical analysis of cosmological topologically massive gravity at the chiral point,” `arXiv:0806.4185 [hep-th]`. Published in ‘Fundamental Interactions: A Memorial Volume for Wolfgang Kummer,’ Editors: Daniel Grumiller, Anton Rebhan and Dimitri Vassilevich, World Scientific, 2010, pp.363-374.

[79] A. Strominger, “A Simple Proof of the Chiral Gravity Conjecture,” `arXiv:0808.0506 [hep-th]`.

[80] S. Carlip, “The Constraint Algebra of Topologically Massive AdS Gravity,” *JHEP* **0810** (2008) 078, `arXiv:0807.4152 [hep-th]`.

[81] W. Li, W. Song, and A. Strominger, “Comment on ’Cosmological Topological Massive Gravitons and Photons’,” `arXiv:0805.3101 [hep-th]`.

[82] Y. S. Myung, “Logarithmic conformal field theory approach to topologically massive gravity,” *Phys.Lett.* **B670** (2008) 220–223, `arXiv:0808.1942 [hep-th]`.

[83] I. Sachs and S. N. Solodukhin, “Quasi-Normal Modes in Topologically Massive Gravity,” *JHEP* **0808** (2008) 003, `arXiv:0806.1788 [hep-th]`.

[84] G. Gibbons, C. Pope, and E. Sezgin, “The General Supersymmetric Solution of Topologically Massive Supergravity,” *Class.Quant.Grav.* **25** (2008) 205005, `arXiv:0807.2613 [hep-th]`.

[85] K. Skenderis, M. Taylor, and B. C. van Rees, “Topologically Massive Gravity and the AdS/CFT Correspondence,” *JHEP* **0909** (2009) 045, `arXiv:0906.4926 [hep-th]`.

[86] A. J. Feingold and I. B. Frenkel, “A hyperbolic Kac-Moody algebra and the theory of Siegel modular forms of genus 2,” *J. Math. Ann.* **263** (1983) 87–144.

[87] V. A. Gritsenko and V. V. Nikulin, “Automorphic forms and Lorentzian Kac-Moody algebras. II,” *Internat. J. Math.* **9** (1998) no. 2, 201–275. `http://dx.doi.org/10.1142/S0129167X98000117`.

[88] V. A. Gritsenko and V. V. Nikulin, “Automorphic forms and Lorentzian Kac-Moody algebras. I,” *Internat. J. Math.* **9** (1998) no. 2, 153–199. `http://dx.doi.org/10.1142/S0129167X98000105`.

[89] V. A. Gritsenko and V. V. Nikulin, “Siegel automorphic form corrections of some Lorentzian Kac-Moody Lie algebras,” *Amer. J. Math.* **119** (1997) no. 1, 181–224. `http://muse.jhu.edu/journals/american_journal_of_mathematics/v119/119.1gritsenko.pdf`.

[90] J. Conway, R. Curtis, S. Norton, R. Parker, and R. Wilson, *Atlas of finite groups. Maximal subgroups and ordinary characters for simple groups. With comput. assist. from J. G. Thackray.* Oxford: Clarendon Press, 1985.

[91] V. Pless, “On the uniqueness of the Golay codes,” *J. Combinatorial Theory* **5** (1968) 215–228.

[92] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups.* Springer-Verlag, 1999.

[93] J. McKay, “Graphs, singularities, and finite groups,” in *The Santa Cruz Conference on Finite Groups (Univ. California, Santa Cruz, Calif., 1979)*, vol. 37 of *Proc. Sympos. Pure Math.*, pp. 183–186. Amer. Math. Soc., Providence, R.I., 1980.

[94] P. Slodowy, *Simple singularities and simple algebraic groups*, vol. 815 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[95] G. Gonzalez-Sprinberg and J.-L. Verdier, “Construction géométrique de la correspondance de McKay,” *Ann. Sci. École Norm. Sup. (4)* **16** (1983) no. 3, 409–449 (1984). `http://www.numdam.org/item?id=ASENS_1983_4_16_3_409_0`.

[96] J. H. Conway, “A simple construction for the Fischer-Griess monster group,” *Invent. Math.* **79** (1985) no. 3, 513–540.

[97] G. Glauberman and S. P. Norton, “On McKay’s connection between the affine E_8 diagram and the Monster,” in *Proceedings on Moonshine and related topics (Montréal, QC, 1999)*, vol. 30 of *CRM Proc. Lecture Notes*, pp. 37–42. Amer. Math. Soc., Providence, RI, 2001.

[98] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.4*, 2005. (`http://www.gap-system.org`).

[99] B. Kostant, “The graph of the truncated icosahedron and the last letter of Galois,” *Notices Amer. Math. Soc.* **42** (1995) no. 9, 959–968.

[100] B. Kostant, “Structure of the truncated icosahedron (e.g. Fullerene or C_{60}, viral coatings) and a 60-element conjugacy class in $\mathrm{PSL}(2, 11)$,” *Selecta Math. (N.S.)* **1** (1995) no. 1, 163–195. `http://dx.doi.org/10.1007/BF01614076`.

[101] C. J. Cummins and J. F. Duncan, “An E_8 Correspondence for Multiplicative Eta-Products,” *Canad. Math. Bull.* **55** (2012) no. 1, 67–72. `http://dx.doi.org/10.4153/CMB-2011-068-9`.

[102] J. F. Duncan, “Arithmetic groups and the affine E_8 Dynkin diagram,” in *Groups and symmetries*, vol. 47 of *CRM Proc. Lecture Notes*, pp. 135–163. Amer. Math. Soc., Providence, RI, 2009.

[103] S. Sakuma, “6-transposition property of τ-involutions of vertex operator algebras,” *Int. Math. Res. Not. IMRN* (2007) no. 9, Art. ID rnm 030, 19. `http://dx.doi.org/10.1093/imrn/rnm030`.

[104] C. H. Lam, H. Yamada, and H. Yamauchi, “Vertex operator algebras, extended E_8 diagram, and McKay’s observation on the Monster simple group,” *Trans. Amer. Math. Soc.* **359** (2007) no. 9, 4107–4123 (electronic). `http://dx.doi.org/10.1090/S0002-9947-07-04002-0`.

[105] C. H. Lam, H. Yamada, and H. Yamauchi, “McKay’s observation and vertex operator algebras generated by two conformal vectors of central charge 1/2,” *IMRP Int. Math. Res. Pap.* (2005) no. 3, 117–181.

[106] G. Höhn, C. H. Lam, and H. Yamauchi, "McKay's E_7 observation on the Baby Monster," *Int. Math. Res. Not. IMRN* (2012) no. 1, 166–212.

[107] G. Höhn, C. H. Lam, and H. Yamauchi, "McKay's E_6 observation on the largest Fischer group," *Comm. Math. Phys.* **310** (2012) no. 2, 329–365. `http://dx.doi.org/10.1007/s00220-011-1413-8`.

[108] T. Eguchi and K. Hikami, "Twisted elliptic genus for $K3$ and Borcherds product," `1112.5928`.

[109] S. Govindarajan, "Brewing moonshine for mathieu," `1012.5732`.

[110] R. Volpato, "Mathieu Moonshine and symmetries of K3 sigma models," `arXiv:1201.6172 [hep-th]`.

[111] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Symmetries of $K3$ sigma models," `1106.4315`.

[112] A. Taormina and K. Wendland, "The overarching finite symmetry group of Kummer surfaces in the Mathieu group M_{24}," `arXiv:1107.3834 [hep-th]`.

[113] B. Gordon and R. J. McIntosh, "A survey of classical mock theta functions," in *Partitions, q-Series, and Modular Forms*, K. Alladi and F. Garvan, eds., vol. 23 of *Developments in Mathematics*, pp. 95–144. Springer New York, 2012. `http://dx.doi.org/10.1007/978-1-4614-0028-8$_9$`. 10.1007/978-1-4614-0028-8_9.

[114] Y. Zhu, "Modular invariance of characters of vertex operator algebras," *Journal of the American Mathematical Society* **9** (1996) no. 1, 237–302.

[115] C. Dong, H. Li, and G. Mason, "Modular invariance of trace functions in orbifold theory and generalized Moonshine," *Communications in Mathematical Physics* **214** (2000) 1–56, `q-alg/9703016`.

[116] M. Benard, "Schur indexes of sporadic simple groups," *J. Algebra* **58** (1979) no. 2, 508–522. `http://dx.doi.org/10.1016/0021-8693(79)90177-7`.

[117] R. S. Margolin, "Representations of M_{12}," *J. Algebra* **156** (1993) no. 2, 362–369. `http://dx.doi.org/10.1006/jabr.1993.1078`.

[118] S. Mukai, "Finite groups of automorphisms of $K3$ surfaces and the Mathieu group," *Invent. Math.* **94** (1988) no. 1, 183–221. `http://dx.doi.org/10.1007/BF01394352`.

[119] S. Kondō, "Niemeier lattices, Mathieu groups, and finite groups of symplectic automorphisms of $K3$ surfaces," *Duke Math. J.* **92** (1998) no. 3, 593–603.

http://dx.doi.org/10.1215/S0012-7094-98-09217-1. With an appendix by Shigeru Mukai.

[120] V. A. Gritsenko and V. V. Nikulin, "$K3$ surfaces, Lorentzian Kac-Moody algebras and mirror symmetry," *Math. Res. Lett.* **3** (1996) no. 2, 211–229.

[121] V. A. Gritsenko and V. V. Nikulin, "The arithmetic mirror symmetry and Calabi-Yau manifolds," *Comm. Math. Phys.* **210** (2000) no. 1, 1–11. http://dx.doi.org/10.1007/s002200050769.

[122] I. V. Dolgachev, "Mirror symmetry for lattice polarized $K3$ surfaces," *J. Math. Sci.* **81** (1996) no. 3, 2599–2630. http://dx.doi.org/10.1007/BF02362332. Algebraic geometry, 4.

[123] R. Dijkgraaf, E. Verlinde, and H. Verlinde, "Counting Dyons in N=4 String Theory," *Nucl.Phys.B* **484:543-561,1997** (Nucl.Phys.B484:543-561,1997) , hep-th/9607026.

[124] A. Sen, "Walls of Marginal Stability and Dyon Spectrum in N=4 Supersymmetric String Theories," *JHEP* **0705:039,2007** (JHEP 0705:039,2007) , hep-th/0702141.

[125] M. C. N. Cheng and E. Verlinde, "Dying Dyons Don't Count," *JHEP* **09** (2007) 070, arXiv:0706.2363 [hep-th].

[126] M. C. Cheng and E. P. Verlinde, "Wall Crossing, Discrete Attractor Flow, and Borcherds Algebra," *SIGMA* **4** (2008) 068, arXiv:0806.2337 [hep-th].

[127] A. Dabholkar, D. Gaiotto, and S. Nampuri, "Comments on the spectrum of CHL dyons," *JHEP* **01** (2008) 023, arXiv:hep-th/0702150.

[128] A. Dabholkar, "Cargese lectures on black holes, dyons, and modular forms," *Nucl. Phys. Proc. Suppl.* **171** (2007) 2–15.

[129] S. Kachru and C. Vafa, "Exact results for N=2 compactifications of heterotic strings," *Nucl.Phys.* **B450** (1995) 69–89, arXiv:hep-th/9505105 [hep-th].

[130] S. Ferrara, J. A. Harvey, A. Strominger, and C. Vafa, "Second quantized mirror symmetry," *Phys.Lett.* **B361** (1995) 59–65, arXiv:hep-th/9505162 [hep-th].

[131] J. A. Harvey and G. W. Moore, "Algebras, BPS states, and strings," *Nucl.Phys.* **B463** (1996) 315–368, arXiv:hep-th/9510182 [hep-th].

[132] R. E. Borcherds, "Automorphic forms with singularities on Grassmannians," *Invent. Math.* **132** (1998) no. 3, 491–562. http://dx.doi.org/10.1007/s002220050232.

[133] M. Kontsevich, “Product formulas for modular forms on $\mathrm{O}(2,n)$ (after R. Borcherds),” *Astérisque* (1997) no. 245, Exp. No. 821, 3, 41–56. Séminaire Bourbaki, Vol. 1996/97.

[134] G. Lopes Cardoso, “Perturbative gravitational couplings and Siegel modular forms in D = 4, N=2 string models,” *Nucl.Phys.Proc.Suppl.* **56B** (1997) 94–101, `arXiv:hep-th/9612200 [hep-th]`.

[135] G. Lopes Cardoso, G. Curio, and D. Lust, “Perturbative couplings and modular forms in N=2 string models with a Wilson line,” *Nucl.Phys.* **B491** (1997) 147–183, `arXiv:hep-th/9608154 [hep-th]`.

[136] S. Govindarajan, “BKM Lie superalgebras from counting twisted CHL dyons,” *JHEP* **1105** (2011) 089, `arXiv:1006.3472 [hep-th]`. Dedicated to the memories of Jaydeep Majumder and Alok Kumar.

[137] `http://modi.countnumber.de/index.php?chap=ell.newforms/ell.newforms.html`.

[138] C. Itzykson, ed., *From Number Theory to Physics*. Springer, 1992.

[139] H. Aoki and T. Ibukiyama, “Simple graded rings of Siegel modular forms, differential operators and Borcherds products,” *Int. J. Math.* **16, No. 3** (2005) 249–279.

Preprint typeset in JHEP style - HYPER VERSION

arXiv:1008.3778v3 [hep-th] 8 Jun 2011

Mathieu Moonshine in the elliptic genus of K3

Matthias R. Gaberdiel, Stefan Hohenegger and Roberto Volpato

Institut für Theoretische Physik
ETH Zurich
CH-8093 Zürich
Switzerland

Abstract: It has recently been conjectured that the elliptic genus of K3 can be written in terms of dimensions of Mathieu group $\mathbb{M}_{24}$ representations. Some further evidence for this idea was subsequently found by studying the twining genera that are obtained from the elliptic genus upon replacing dimensions of Mathieu group representations by their characters. In this paper we find explicit formulae for all (remaining) twining genera by making an educated guess for their general modular properties. This allows us to identify the decomposition of *all* expansion coefficients in terms of dimensions of $\mathbb{M}_{24}$-representations. For the first 500 coefficients we verify that the multiplicities with which these representations appear are indeed all non-negative integers. This represents very compelling evidence in favour of the conjecture.

Contents

1. **Introduction** 1

2. **Elliptic genus and twining characters** 3

3. **Modular properties of the twining genera** 5
   3.1 Characters evaluated at fixed values of z 8
   3.2 Studying the modular forms of weight two 10

4. **Decomposition into irreducible representations** 11

5. **Conclusions** 12

A. **Definitions** 16

B. **Modular forms for $\Gamma_0(N)$** 17

C. **The modular forms of weight two** 19

1. Introduction

A few months ago Eguchi, Ooguri and Tachikawa [1] observed that the elliptic genus of K3 seems to involve representations of the largest of the Mathieu groups, $\mathbb{M}_{24}$. More specifically, they studied the expansion of the elliptic genus of K3 in terms of elliptic genera of $\mathcal{N}=4$ superconformal representations following [2], and noted that the expansion coefficients can be written in terms of dimensions of representations of $\mathbb{M}_{24}$. This intriguing observation is very reminiscent of a similar phenomenon usually referred to as 'Monstrous Moonshine', namely that the famous J-function has an expansion in terms of characters of the Virasoro algebra whose coefficients are dimensions of Monster group representations, as was first noted by McKay and Thompson.

In the context of Monstrous Moonshine, this observation was eventually explained by the construction of the so-called Monster conformal field theory $V^\natural$ [3], a self-dual conformal field theory at $c=24$ whose space of states is of the form

$$V^\natural = \bigoplus_{n=0}^{\infty} \left(V_n \otimes \mathcal{H}^{\text{Vir}}_{h=n} \right) \ , \tag{1.1}$$

where each V_n is a representation of the Monster group, while $\mathcal{H}^{\text{Vir}}_h$ denotes the irreducible Virasoro representation with conformal weight h and $c=24$. The J-function is then the partition function of $V^\natural$, and its Fourier coefficients are therefore sums of dimensions of irreducible Monster group representations. Furthermore, the automorphism group of $V^\natural$ is the Monster group.

One key observation that provided convincing evidence for (1.1) came from considering the so-called McKay-Thompson series [4]. These are obtained from the J-function upon replacing the expansion coefficients $A_n = \dim(V_n)$ by their corresponding characters, $\mathrm{Tr}_{V_n}(g)$, where g is an element of the Monster group.[1] It was shown by Conway & Norton [5] that these McKay-Thompson series have nice modular properties under congruence subgroups of $SL(2,\mathbb{Z})$. This is what one expects if they arise indeed from (1.1) since they are then equal to the 'twining character', *i.e.* the character with the insertion of the group element g, which has good modular properties based on standard orbifold arguments. For a review of these and other aspects of Monstrous Moonshine see *e.g.* [6].

By analogy to (1.1) the observation of [1] suggests that the states that contribute to the elliptic genus of K3 have the structure

$$\mathcal{H}^{\mathrm{BPS}} = \bigoplus_n \left(H_n \otimes \mathcal{H}_n^{\mathcal{N}=4} \right) \ , \tag{1.2}$$

where the sum runs over all irreducible $\mathcal{N} = 4$ representations that contribute to the elliptic genus, while each H_n is a $\mathbb{M}_{24}$ representation. In order to test this idea it is then again natural to consider the 'twining genera' $\phi_g(\tau, z)$, where one replaces $\dim(H_n)$ by the corresponding character $\mathrm{Tr}_{H_n}(g)$, with $g \in \mathbb{M}_{24}$ [7, 8]. Unfortunately, the dimensions of the irreducible representations of $\mathbb{M}_{24}$ are rather small, and the decomposition of the expansion coefficients of the elliptic genus of K3 in terms of dimensions of $\mathbb{M}_{24}$ representations could only be guessed reliably for the first few coefficients. However, the elliptic genus of K3 is a weak Jacobi form of index one and weight zero [9], and one expects that the twining genera should have similar properties. More specifically, if $g \in \mathbb{M}_{24}$ has order N, then the corresponding twining genus $\phi_g(\tau, z)$ should transform as a weak Jacobi form of index 1 and weight 0 under the congruence subgroup $\Gamma_0(N)$, possibly up to a multiplier system.

In [7, 8] this assumption about the modular properties of the twining genera, together with the explicit knowledge of the first few coefficients, was used in order to determine some of them explicitly. The fact that these two constraints were compatible was already a fairly non-trivial consistency check on the proposal of [1]. Furthermore, for group elements that enjoy an interpretation as an automorphism symmetry of K3 at a suitable point in moduli space, one could calculate the corresponding twining genera directly (see in particular [10]). In this way some of these formulae could be confirmed independently [7].

In this paper we shall complete the analysis of [7, 8] by finding explicit formulae for the remaining conjugacy classes.[2] In particular, we shall make a precise proposal for the structure of the multiplier system that appears in the modular transformation formula, see (3.8) and table 1 below. We can then follow again the strategies outlined in [8] and [7], respectively, to determine the twining genera, using the explicit knowledge of the first few terms. The fact that this procedure is successful is again a fairly non-trivial consistency check of the proposed structure (1.2).

---

[1] These series can be calculated provided one knows how to write the Fourier coefficients of the J-function as $A_n = \dim(V_n)$, but do not require any additional knowledge.

[2] After posting the first version of this paper on the arXiv, we were informed by Tohru Eguchi that they have independently obtained explicit formulae for these twining genera which agree with ours, see [24].

The knowledge of *all* twining genera also leads to another, highly constraining, consistency check. If (1.2) holds and if, for each g, our explicit formula for ϕ_g really corresponds to the trace over $\mathcal{H}^{\rm BPS}$ with the insertion of g, then its coefficients must be equal to $\mathrm{Tr}_{H_n}(g)$ for all n. But knowing these traces for all classes g is sufficient to identify unambiguously H_n as a (possibly reducible) representation of $\mathbb{M}_{24}$. Thus assuming that (1.2) holds, we can *deduce* the decomposition of H_n into irreducible $\mathbb{M}_{24}$ representations for all n. We have calculated the multiplicities explicitly up to order 500 — the results up to order 30 are tabulated in table 4 — and they are all indeed non-negative integers. The fact that this decomposition works out, *i.e.* that each H_n can be written as a direct sum of irreducible representations (with integer multiplicities) is then a highly non-trivial check of (1.2); in fact, we would in some sense prove the proposal of [1] if we could show that *all* of these multiplicities are indeed non-negative integers.

The paper is organised as follows. In section 2 we briefly review some of the properties of the elliptic genus of K3, as well as the proposal of [1]. The modular properties of the twining genera are explained in detail in section 3. In particular, by analogy with the situation for the McKay-Thompson series of Monstrous Moonshine, we make a specific proposal for the structure of the multiplier system. This proposal is then subsequently verified in sections 3.1 and 3.2. In section 3.1 we adopt the strategy of [8] and consider the twining genus at specific values of z, finding explicit formulae for all remaining cases. Section 3.2, on the other hand, employs the method of [7] to write the twining genus in terms of a modular form of weight two; this allows us to check some of the modular properties of the twining genera fairly directly, but the analysis is quite complicated and the details are described in appendix B and C. Finally, we explain in section 4 how the knowledge of all twining genera determines the decomposition of the coefficients in terms of $\mathbb{M}_{24}$ representations, and give the explicit results in table 4. We close with some comments and speculations in section 5.

2. Elliptic genus and twining characters

The elliptic genus of an $\mathcal{N}=2$ superconformal algebra is defined by

$$\phi(\tau,z)=\mathrm{Tr}_{\mathcal{H}_{\rm RR}}\left(q^{L_0-\frac{c}{24}}e^{2\pi i z J_0}(-1)^F\,\bar{q}^{\bar{L}_0-\frac{\bar{c}}{24}}(-1)^{\bar{F}}\right)\,, \tag{2.1}$$

where $q=e^{2\pi i\tau}$, and the trace is taken in the RR sector. For the right-movers (whose modes are denoted by a bar), only the ground states contribute, and hence the above expression is in fact independent of $\bar{q}$. As is well known [9], the modularity properties of conformal field theory together with spectral flow invariance and unitarity imply that the elliptic genus is a *weak Jacobi form* of index $m=\frac{c}{6}$ and weight 0 [11]. A weak Jacobi form $\phi(\tau,z)$ of weight w and index $m\in\mathbb{Z}$ is a function ϕ of $(\tau,z)\in\mathbb{H}\times\mathbb{C}$, where $\mathbb{H}$ is the upper half-plane. It is characterised by the transformation properties

$$\phi\Big(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\Big) = (c\tau+d)^w\, e^{2\pi i m \frac{cz^2}{c\tau+d}}\,\phi(\tau,z) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})\ , \tag{2.2}$$

$$\phi(\tau, z+\ell\tau+\ell') = e^{-2\pi i m(\ell^2\tau+2\ell z)}\phi(\tau,z) \qquad \ell,\ell'\in\mathbb{Z}\ , \tag{2.3}$$

and has a Fourier expansion

$$\phi(\tau,z) = \sum_{n\geq 0,\ell\in\mathbb{Z}} c(n,\ell)q^n y^\ell \tag{2.4}$$

where $y = e^{2\pi i z}$ and $c(n,\ell) = (-1)^w c(n,-\ell)$. For the case of K3 that will concern us primarily in this paper, $m = 1$ and the elliptic genus equals [12]

$$\phi_{K3}(\tau,z) = 2y + 20 + 2y^{-1} + q\Big(20y^2 - 128y + 216 - 128y^{-1} + 20y^{-2}\Big) + \mathcal{O}(q^2)\ . \tag{2.5}$$

It can be thought of as the partition function of the $\mathcal{N}=2$ half-BPS states of type II string theory on K3.

For the case of K3 the conformal field theory is actually $\mathcal{N}=4$ superconformal, and one can therefore write the elliptic genus in terms of the elliptic genera associated to $\mathcal{N}=4$ superconformal representations. It was observed in [1], following on from earlier work [2], that it can be written as

$$\phi_{K3}(\tau,z) = 24\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac14,l=0}(\tau,z) + \sum_{n=0}^{\infty} A_n\,\mathrm{ch}^{\mathcal{N}=4}_{h=n+\frac14,l=\frac12}(\tau,z)\ , \tag{2.6}$$

where $\mathrm{ch}^{\mathcal{N}=4}_{h=\frac14,l=0}$ is the elliptic genus of the short $\mathcal{N}=4$ representation with $h=\frac14$ and $l=0$ — see [13, 14] for an explicit formula — while

$$\mathrm{ch}^{\mathcal{N}=4}_{h,l=\frac12}(\tau,z) = q^{h-\frac38}\,\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^3} \tag{2.7}$$

is the elliptic genus of a long $\mathcal{N}=4$ representation.[3] The observation of [1] was that the coefficients A_n can be written in terms of dimensions of representations H_n of the Mathieu group $\mathbb{M}_{24}$, so that

$$\begin{aligned}\phi_{K3}(\tau,z) =& (\dim H_{00})\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac14,l=0}(\tau,z) - (\dim H_0)\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac14,l=\frac12}(\tau,z)\\ &+\sum_{n=1}^{\infty}(\dim H_n)\,\mathrm{ch}^{\mathcal{N}=4}_{h=n+\frac14,l=\frac12}(\tau,z)\ ,\end{aligned} \tag{2.8}$$

where

$$\begin{aligned} H_{00} &= \mathbf{23}+\mathbf{1} & H_0 &= 2\cdot\mathbf{1}\\ H_1 &= \mathbf{45}+\overline{\mathbf{45}} & H_2 &= \mathbf{231}+\overline{\mathbf{231}}\\ H_3 &= \mathbf{770}+\overline{\mathbf{770}} & H_4 &= \mathbf{2277}+\overline{\mathbf{2277}}\\ H_5 &= 2\cdot\mathbf{5796} & H_6 &= 2\cdot\mathbf{3520}+2\cdot\mathbf{10395}\end{aligned} \tag{2.9}$$

---

[3] Strictly speaking, the $\mathcal{N}=4$ representation with $n=0$ ($h=\frac14$) is short, and thus (2.7) for $h=\frac14$ is not the elliptic genus of a single representation, but rather involves a sum of representations.

$$H_7 = 2\cdot\mathbf{1771} + 2\cdot\mathbf{2024} + 2\cdot\mathbf{5313} + 2\cdot\mathbf{5796} + 2\cdot\mathbf{5544} + 2\cdot\mathbf{10395}\ . \tag{2.10}$$

The dimensions of the irreducible representations of $\mathbb{M}_{24}$ can be read off from the character table (see table 3). Note that we have absorbed the prefactor 2 in equation (1.11) of [1] into the definition of $A_n = \dim(H_n)$. Then we can write the H_n in terms of real representations, so that, for example, H_1 is the sum of a pair of conjugate representations. The expression for H_7 was given in [7, 8] and differs from what was originally proposed in [1].

It is natural to conjecture that such a decomposition is the hallmark of a deeper structure underlying the elliptic genus of K3, see (1.2). In order to test this idea, the '*twining elliptic genera*', *i.e.* the analogues of the McKay-Thompson series of Monstrous Moonshine, were considered in [7, 8]. These twining genera are obtained from the elliptic genus upon inserting a group element $g \in \mathbb{M}_{24}$ into the trace

$$\phi_g(\tau, z) = \frac{1}{2}\,\mathrm{Tr}_{\mathcal{H}_{\mathrm{RR}}}\Big(g\, q^{L_0 - \frac{c}{24}} e^{2\pi i z J_0} (-1)^F\, \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} (-1)^{\bar{F}}\Big)\ . \tag{2.11}$$

As in [8], we shall normalise the twining characters so that $\phi_{1\mathrm{A}}(\tau,z) = \frac{1}{2}\phi_{K3}(\tau,z)$ is directly equal to the standard weak Jacobi form $\phi_{0,1}$ (see appendix A). Technically speaking, ϕ_g is simply obtained from (2.8) by replacing the dimensions $A_n = \dim(H_n)$ by the trace of g over H_n, $A_n(g) = \mathrm{Tr}_{H_n}(g)$, *i.e.*

$$\begin{aligned}\phi_g(\tau, z) = \frac{1}{2}\Big[&\mathrm{Tr}_{H_{00}}(g)\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=0}(\tau,z) - \mathrm{Tr}_{H_0}(g)\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=\frac{1}{2}}(\tau,z)\\ &+ \sum_{n=1}^{\infty}\mathrm{Tr}_{H_n}(g)\,\mathrm{ch}^{\mathcal{N}=4}_{h=n+\frac{1}{4},l=\frac{1}{2}}(\tau,z)\Big]\ .\end{aligned} \tag{2.12}$$

The character of g only depends on its conjugacy class, and thus the various traces can be read off from the character table of $\mathbb{M}_{24}$, see table 3.

As discussed in more detail in the next section, the twining genera are expected to be Jacobi forms under suitable congruence subgroups of $\mathrm{SL}(2,\mathbb{Z})$. This was confirmed for a number of conjugacy classes in [7, 8]. In the next section we shall complete this programme by determining the twining characters for all remaining conjugacy classes. We shall furthermore show that they have the appropriate modular properties.

3. Modular properties of the twining genera

Using standard conformal field theory arguments, it was argued in [7, 8] that the twining genera ϕ_g should transform as Jacobi forms of index 1 and weight 0, possibly up to a phase, under the congruence subgroup $\Gamma_0(N)$. More specifically, this means that ϕ_g satisfies (2.3) for all $\ell, \ell' \in \mathbb{Z}$, while (2.2) only holds for

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \mid c \equiv 0 \mod N \right\}\ , \tag{3.1}$$

where N is the order of g. In [7], this was explicitly verified for the conjugacy classes[4]

$$1\mathrm{A},\ 2\mathrm{A},\ 3\mathrm{A},\ 4\mathrm{B},\ 5\mathrm{A},\ 6\mathrm{A},\ 7\mathrm{AB},\ 8\mathrm{A},\ 11\mathrm{A},\ 14\mathrm{AB},\ 15\mathrm{AB},\ 23\mathrm{AB}\ . \tag{3.2}$$

These classes are characterised by the condition $\phi_g(\tau, 0) \neq 0$. This is equivalent to the condition that a representative of the class is contained in the subgroup $\mathbb{M}_{23} \subset \mathbb{M}_{24}$, where we think of $\mathbb{M}_{24}$ as a subgroup of S_{24}, the permutation group of 24 points, and define $\mathbb{M}_{23} \subset \mathbb{M}_{24}$ to be the subgroup fixing, say, the first point. All geometric symmetries of K3 at a suitable point in moduli space lie in this subgroup [15, 16], and thus some of them can be calculated from first principles, see also [10]. For all of them the multiplier system turned out to be trivial [7]. The situation is more difficult for the remaining conjugacy classes

$$2\mathrm{B},\ 3\mathrm{B},\ 4\mathrm{A},\ 4\mathrm{C},\ 6\mathrm{B},\ 10\mathrm{A},\ 12\mathrm{A},\ 12\mathrm{B},\ 21\mathrm{AB}\ , \tag{3.3}$$

since there is no *a priori* method to determine them. In [8], explicit formulae were found for the first few of them by combining the constraints from modularity with the knowledge of the first few coefficients.[5] In fact, the analysis of [8] was performed for the NS-sector version of the twining genus, the twining character

$$\chi_g(\tau, z) = \exp\left[2\pi i\left(\frac{\tau}{4} + z + \frac{1}{2}\right)\right]\phi_g\left(\tau, z + \frac{\tau}{2} + \frac{1}{2}\right) \tag{3.4}$$

evaluated at $z = 0$. The advantage of this approach is that one can work with standard modular functions (rather than Jacobi forms). The price one has to pay, on the other hand, is that part of the modular invariance is broken, and that multiplier phases are introduced for certain modular transformations. The latter property turned out to be a blessing in disguise since it suggested that multiplier phases may naturally appear in the modular transformations. Indeed, in [8] the twining characters were determined for all elements g up to order $o(g) \leq 6$, and it was found that the classes in (3.2) and in (3.3) appear to behave very similarly. This suggests that also the twining genera ϕ_g associated to (3.3) should be invariant under $\Gamma_0(N)$, possibly up to non-trivial phases.

The appearance of a multiplier system in the transformation rule for the twining genera is certainly consistent with standard CFT arguments. In fact, this phenomenon also occurs for several McKay-Thompson series (and, more generally, for replicable functions [17]), which are the analogues of the twining characters in the context of Monstrous Moonshine [5]. Recall that each McKay-Thompson series T_g is associated with a certain discrete group

$$\Gamma_0(N|h) = \left\{\begin{pmatrix} a & b/h \\ Nc & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}) \mid a,b,c,d \in \mathbb{Z}\right\}\ , \tag{3.5}$$

where N is the order of the Monster class g, and h is some integer such that $h|\gcd(N, 24)$. This group (sometimes with the inclusion of some Atkin-Lehner involutions of $\Gamma_0(N)$, see

---

[4] The classes which are power conjugated, for example 7A and 7B, give rise to the same twining genus, so that we denote them as a unique class 7AB. The twining genera for the classes 1A, 2A, 3A, 4B, 5A and 6A were also found in [8].

[5] The twining characters for 2B and 4A were also found in [7].

[5]) is the (restricted) eigengroup of T_g, *i.e.* the group under which T_g is invariant up to h-th roots of unity. In particular, T_g is invariant (without any phases) under $\Gamma_0(Nh) \subset \Gamma_0(N)$, while under the cosets of $\Gamma_0(N|h)/\Gamma_0(Nh)$, that are represented by

$$\begin{pmatrix} 1 & 1/h \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix} , \tag{3.6}$$

it transforms as

$$T_g\Big(\tau + \frac{1}{h}\Big) = e^{-\frac{2\pi i}{h}} T_g(\tau) \qquad \text{and} \qquad T_g\Big(\frac{\tau}{N\tau+1}\Big) = e^{\pm\frac{2\pi i}{h}} T_g(\tau) . \tag{3.7}$$

The two cosets in (3.6) generate $\Gamma_0(N|h)$, so that (3.7) uniquely determines the multiplier system under $\Gamma_0(N|h)$.

It is then natural to expect that analogous properties hold for the twining genera of $\mathbb{M}_{24}$. The most obvious generalisation would be to require the twining genus ϕ_g, with $g \in \mathbb{M}_{24}$ and $o(g) = N$, to be a Jacobi form (with a suitable multiplier system) of weight 0 and index 1 under $\Gamma_0(N|h)$, for some $h|\gcd(N, 24)$. However, there is one immediate problem with this proposal: for $h > 1$, $\Gamma_0(N|h)$ is not contained in $\mathrm{SL}(2,\mathbb{Z})$, and it is not clear how to define the action of the whole $\Gamma_0(N|h)$ on Jacobi forms.[6] Thus we can only analyse the modular properties under the subgroup $\Gamma_0(N|h) \cap \mathrm{SL}(2,\mathbb{Z}) \cong \Gamma_0(N)$. This then leads to the following conjecture:

Conjecture. *For all the conjugacy classes g of $\mathbb{M}_{24}$, the twining character $\phi_g(\tau, z)$ is a Jacobi form of index one and weight zero under $\Gamma_0(N)$, with a multiplier system defined by*[7]

$$\phi_g\Big(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\Big) = e^{\frac{2\pi i cd}{Nh}}\, e^{\frac{2\pi i\, cz^2}{c\tau+d}}\, \phi_g(\tau, z) , \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) , \tag{3.8}$$

where N is the order of g and $h|\gcd(N, 12)$. The multiplier system is trivial ($h = 1$) if and only if g contains a representative in $\mathbb{M}_{23} \subset \mathbb{M}_{24}$.

For the classes in (3.2) that have representatives in $\mathbb{M}_{23}$ the conjecture has been shown in [7]. In the next subsections, we will show that the conjecture is also true for the remaining classes, *i.e.* the elements in (3.3), with the values of h as given in table 1.

Class	2B	3B	4A	4C	6B	10A	12A	12B	21AB
h	2	3	2	4	6	2	2	12	3

Table 1: Value of h for the conjugacy classes in (3.3).

Because of these non-trivial multiplier systems, the analysis is quite difficult, and we have applied two different strategies. First we have refined the method of [8] by considering the twining genus $\phi_g(\tau, z)$ as a function of τ at special values of z. These values are chosen in such a way that the z-dependent exponential factor in the transformation formula (3.8) cancels (part of) the h-th root of unity. If the phase is completely removed (as is the case for all but two classes), the resulting function is a modular function for $\Gamma_0(N)$, which can be easily analysed. Using this approach (as well as some guess work for the other two cases) we have succeeded in finding closed formulae for all the characters with $h > 1$, see section 3.1.

---

[6] A well-defined action of $\mathrm{SL}(2,\mathbb{R})$ on Jacobi forms can be defined, see [11]. However, this action does not respect the periodicity condition on z, and thus does not seem to be relevant in the current context.

[7] We thank Miranda Cheng and John Duncan for pointing out an error in a previous version of this formula.

The other strategy follows the idea advocated in [7] and consists of expanding the twining genus in terms of standard weak Jacobi forms. This reduces the problem to finding a suitable modular form of weight two. This problem can be studied systematically, but usually leads to more complicated computations. In section 3.2 we shall explain the salient features of this analysis, while the explicit formulae for all characters are given in appendix C.

3.1 Characters evaluated at fixed values of z

Let $\phi(\tau, z)$ be a weak Jacobi form of weight 0 and index 1, transforming as in (3.8), for some N and h with $h|\gcd(N, 12)$. Then, for any $k \in \mathbb{Z}$, we have

$$\begin{aligned}
\phi\Big(\frac{a\tau+b}{c\tau+d}, \frac{k}{h}\Big) &= \phi\Big(\frac{a\tau+b}{c\tau+d}, \frac{k(c\tau+d)}{h(c\tau+d)}\Big) = e^{\frac{2\pi i cd}{Nh}}\, e^{\frac{2\pi i\, k^2 c(c\tau+d)}{h^2}}\, \phi\Big(\tau, \frac{k(c\tau+d)}{h}\Big) \\
&= e^{\frac{2\pi i cd}{Nh}}\, e^{\frac{2\pi i k^2 c(c\tau+d)}{h^2}}\, e^{-2\pi i(\frac{k^2c^2}{h^2}\tau + 2\frac{k^2 cd}{h^2})}\, \phi\Big(\tau, \frac{kd}{h}\Big) \qquad (3.9)\\
&= e^{2\pi i cd(-\frac{k^2}{h^2} + \frac{1}{Nh})}\, \phi\Big(\tau, \frac{kd}{h}\Big)\ , \qquad \text{for} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)\ ,
\end{aligned}$$

where we have used (2.3) in the second line. (Note that $c \in N\mathbb{Z}$, and hence $\frac{c}{h} \in \mathbb{Z}$.) Let us define

$$\Phi^{(h)}(\tau) = \frac{1}{\varphi(h)} \sum_{k \in (\mathbb{Z}/h\mathbb{Z})^*} \phi\Big(\tau, \frac{k}{h}\Big)\ , \qquad (3.10)$$

where $(\mathbb{Z}/h\mathbb{Z})^*$ is the set of totatives of h, *i.e.* the positive integers smaller than h that are relatively prime to h, and the Euler totient function $\varphi(h)$ is the number of all such totatives. This definition simplifies considerably in concrete examples

$$\Phi^{(h)}(\tau) = \phi(\tau, \tfrac{1}{h})\ , \qquad h = 2, 3, 4, 6\ , \qquad (3.11)$$

$$\Phi^{(12)}(\tau) = \frac{1}{2}\big(\phi(\tau, \tfrac{1}{12}) + \phi(\tau, \tfrac{5}{12})\big)\ , \qquad (3.12)$$

because $\phi(\tau, z) = \phi(\tau, -z)$ for Jacobi forms of even weight. It is easy to verify that, for all $h|12$, the condition $\gcd(k, h) = 1$ implies $k^2 \equiv 1 \mod h$. Furthermore, for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$, the condition $ad - bc = 1$ implies $\gcd(d, h) = 1$, so that the map $k \mapsto kd$ is bijective on $(\mathbb{Z}/h\mathbb{Z})^*$. Thus we conclude

$$\Phi^{(h)}\Big(\frac{a\tau+b}{c\tau+d}\Big) = e^{2\pi i cd(-\frac{1}{h^2} + \frac{1}{Nh})}\, \Phi^{(h)}(\tau)\ , \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)\ , \qquad (3.13)$$

and it follows immediately that $\Phi^{(h)}(\tau)$ is invariant under $\Gamma_0(Nh)$. Since h divides N, the cosets of $\Gamma_0(N)/\Gamma_0(Nh)$ are generated by $\left(\begin{smallmatrix} 1 & 0 \\ N & 1 \end{smallmatrix}\right)$, and the phase in (3.13) cancels under this transformation if and only if

$$\frac{N}{h} \equiv 1 \mod h \ . \tag{3.14}$$

Thus $\Phi^{(h)}$ is actually invariant under $\Gamma_0(N)$ if (3.14) holds. In this case, it is easy to obtain a closed formula for $\Phi_g^{(h)}$. In fact, with the exception of $N = 21$, all the groups $\Gamma_0(N)$ we are interested in are genus zero[8], so that all modular functions must be rational functions of the corresponding Hauptmodul.

As it turns out the condition (3.14) is satisfied for all classes in (3.3), with the exception of 4A and 12A for which $h = 2$ and hence $N/h \equiv 0 \mod h$. In all other cases we found a modular function for $\Gamma_0(N)$ which matches the first few coefficients in the q-expansion that can be determined from (2.9) and (2.10). This function is either a constant or a fractional linear transformation of the Hauptmodul (in fact, a McKay-Thompson series) for $\Gamma_0(N)$. For the class 21AB it is a rational function in the McKay-Thompson series $T_{[21\mathrm{B}]}$ and $T_{[21\mathrm{D}]}$, that are modular functions for $\Gamma_0(21)$ (for a q-expansion of these McKay Thompson series see appendix A). Our explicit expressions are:

$$\begin{aligned}
\Phi^{(2)}_{2\mathrm{B}}(\tau) &= -4 && h=2\\
\Phi^{(3)}_{3\mathrm{B}}(\tau) &= -3 && h=3\\
\Phi^{(4)}_{4\mathrm{C}}(\tau) &= -2 && h=4\\
\Phi^{(6)}_{6\mathrm{B}}(\tau) &= -1 && h=6\\
\Phi^{(2)}_{10\mathrm{A}}(\tau) &= -4\frac{\eta(5\tau)\eta(2\tau)^5}{\eta(10\tau)\eta(\tau)^5} = \frac{-20}{T_{[10\mathrm{E}]}-3} - 4 && h=2\\
\Phi^{(12)}_{12\mathrm{B}}(\tau) &= -2 - 6\frac{\eta(2\tau)^2\eta(3\tau)\eta(12\tau)^3}{\eta(\tau)^3\eta(4\tau)\eta(6\tau)^2} = \frac{-6}{T_{[12\mathrm{I}]}-3} - 2 && h=12\\
\Phi^{(3)}_{21AB}(\tau) &= \frac{1}{2} - \frac{7}{2}\frac{\eta(3\tau)\eta(7\tau)^3}{\eta(\tau)^3\eta(21\tau)} = \frac{-T_{[21\mathrm{B}]}+7T_{[21\mathrm{D}]}+15}{-2T_{[21\mathrm{B}]}+2} && h=3\ .
\end{aligned} \tag{3.15}$$

Since $\phi_g(\tau, z)$ is a Jacobi form of index one, we have

$$\phi_g(\tau,z) = \phi_g(\tau,0)\,\frac{\vartheta_2(\tau,z)^2}{\vartheta_2(\tau,0)^2} + \phi_g(\tau,\tfrac{1}{2})\,\frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2}\ , \tag{3.16}$$

so that (3.15) immediately gives a formula for the corresponding character at generic z

$$\phi_g(\tau,z) = \Phi_g^{(h)}(\tau)\,\varphi(h)\left(\sum_{k\in(\mathbb{Z}/h\mathbb{Z})^*}\vartheta_1\Big(\tau,\frac{k}{h}\Big)^2\right)^{-1}\vartheta_1(\tau,z)^2\ , \tag{3.17}$$

where we also used the fact that, for all g in (3.3),

$$\phi_g(\tau,0) = \frac{1}{2}\,\mathrm{Tr}_{\mathbf{23}\oplus\mathbf{1}}(g) = 0\ . \tag{3.18}$$

---

[8] Genus zero here means that the Riemann surface obtained by quotienting the upper half-plane $\mathbb{H}$ by $\Gamma_0(N)$ has the topology of the sphere.

The remaining two cases, 4A and 12A, are not invariant under $\Gamma_0(N)$, and thus require more work. For 4A a closed formula was already found in [7, 8]

$$\phi_{4\mathrm{A}}(\tau,\tfrac{1}{2}) = -4 - 32\frac{\eta(2\tau)^2\eta(8\tau)^4}{\eta(\tau)^4\eta(4\tau)^2} = -4 - \frac{32}{T_{8E}(\tau)+4} \qquad h=2\ , \tag{3.19}$$

and it is easy to verify that the corresponding $\phi_{4\mathrm{A}}(\tau,z)$ transforms as in (3.8) under $\Gamma_0(4)$ with $h=2$. We have also managed to find a closed formula for the NS-character χ_g of 12A at $z=\frac{1}{6}$

$$\chi_{12\mathrm{A}}(\tau,\tfrac{1}{6}) = \frac{\eta(\tau)\eta(\frac{3\tau}{2})^2\eta(4\tau)^2\eta(6\tau)^3}{\eta(2\tau)\eta(\frac{\tau}{2})^2\eta(12\tau)^2\eta(3\tau)^3}\ , \tag{3.20}$$

from which the function $\phi_{12\mathrm{A}}(\tau,z)$ can be reconstructed. However, it is easier to analyse the modular properties of $\phi_{12\mathrm{A}}(\tau,z)$ using the methods described in the next subsection.

3.2 Studying the modular forms of weight two

Every weak Jacobi form of index 1 can be written as a linear combination of the standard Jacobi forms $\phi_{0,1}$ and $\phi_{-2,1}$ of weight 0 and -2, respectively (see appendix A), where the coefficients lie in the space of modular forms under the relevant subgroup of $\mathrm{SL}(2,\mathbb{Z})$ [11]. In particular,

$$\phi_g(\tau,z) = B_g\,\phi_{0,1}(\tau,z) + F_g(\tau)\,\phi_{-2,1}(\tau,z)\ , \tag{3.21}$$

where B_g is the constant

$$B_g = \frac{1}{12}\phi_g(\tau,0) = \frac{1}{24}\,\mathrm{Tr}_{\mathbf{23}\oplus\mathbf{1}}(g)\ , \tag{3.22}$$

while F_g is a suitable modular form of weight two. If g is in (3.2), the phases in (3.8) are trivial, and $F_g \in M_2(\Gamma_0(N))$, where $M_k(\Gamma)$ denotes the space of modular form of weight k under the group Γ. In [7], it has been shown that, for all g in (3.2), there is a unique $F_g \in M_2(\Gamma_0(N))$ matching the known coefficients of the q-expansion of ϕ_g (see also appendix C).

If g is one of the classes in (3.3), then $B_g=0$ and the conjecture (3.8) would imply that F_g transforms as

$$F_g\Big(\frac{a\tau+b}{c\tau+d}\Big) = e^{\frac{2\pi icd}{Nh}}(c\tau+d)^2F(\tau)\ , \qquad \begin{pmatrix} a & b\\ c & d\end{pmatrix} \in \Gamma_0(N)\ , \tag{3.23}$$

where N is the order of g and $h|\gcd(N,12)$. In particular, this means that

$$F_g \in M_2(\Gamma_0(Nh))\ , \qquad F_g^h \in M_{2h}(\Gamma_0(N))\ . \tag{3.24}$$

Conversely, if these conditions hold, then F_g transforms as a modular form of weight 2 under $\Gamma_0(N)$ up to a certain h-th root of unity $e^{\frac{2\pi ircd}{Nh}}$, for some $r\in\mathbb{Z}$.

For all g in (3.3), we have found modular forms $F_g \in M_2(\Gamma_0(Nh))$ for the above values of h that reproduce the first few coefficients of q as determined from (2.9) and (2.10), and satisfy $F_g^h \in M_{2h}(\Gamma_0(N))$. The explicit expressions for F_g (and F_g^h) in terms of suitable bases of $M_2(\Gamma_0(Nh))$ (and $M_{2h}(\Gamma_0(N))$) are quite complicated, and are therefore only given in appendix C.

It remains to prove that the h-th root of unity $e^{\frac{2\pi ircd}{Nh}}$ in the modular transformation of F_g is the same as in (3.8), *i.e.* that $r = 1$. We are mainly interested in the class 12A because the other cases have already been dealt with in the previous subsection. For 12A we have found $h = 2$, and since the phase is non-trivial (as can be easily checked), the only possibility is $r = 1$. This completes our analysis.

4. Decomposition into irreducible representations

In the previous sections, we provided explicit expressions for all twining genera, matching the first few coefficients in (2.12) and transforming as in (3.8). The knowledge of all such characters also leads to another very stringent consistency check of the proposal. It allows us to check whether there are indeed underlying $\mathbb{M}_{24}$ representations H_n such that $A_n(g) = \mathrm{Tr}_{H_n}(g)$, where $A_n(g)$ is the coefficient replacing A_n in ϕ_g, see (2.12).

In order to understand how this works, let us assume that $A_n(g) = \mathrm{Tr}_{H_n}(g)$, where H_n is a $\mathbb{M}_{24}$ representation which we decompose as

$$H_n = \bigoplus_{i=1}^{26} h_{n,i}\, R_i \ , \tag{4.1}$$

where $i = 1, \ldots, 26$ labels the different irreducible representations of $\mathbb{M}_{24}$, numbered as in table 3. To start with we rewrite (2.12) as

$$-\mathrm{Tr}_{H_0}(g) + \sum_{n=1}^{\infty} \mathrm{Tr}_{H_n}(g) q^n = q^{\frac{1}{8}} \frac{\eta(\tau)^3}{\vartheta_1(\tau,z)^2} \Big(2\phi_g(\tau,z) - \mathrm{Tr}_{H_{00}}(g)\, \mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=0}(\tau,z) \Big) \ , \tag{4.2}$$

where we have used (2.7). Next we recall that the characters of a finite group satisfy the orthonormality relations

$$\sum_g c(g)\ \overline{\mathrm{Tr}_R(g)}\ \mathrm{Tr}_{R'}(g) = \begin{cases} 1 & \text{if } R \cong R' \\ 0 & \text{otherwise,} \end{cases} \tag{4.3}$$

where R and R' are two irreducible representations of $\mathbb{M}_{24}$. Here the sum runs over all conjugacy classes of $\mathbb{M}_{24}$, and $c(g)^{-1}$ is the order of the centraliser of g

$$c(g) = \frac{n(g)}{|\mathbb{M}_{24}|} \ , \qquad \text{with} \quad |\mathbb{M}_{24}| = 2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 \tag{4.4}$$

the order of $\mathbb{M}_{24}$ and $n(g)$ the number of elements in the conjugacy class of g. Using (4.3) we now obtain from (4.2)

$$-2\delta_{i,1} + \sum_{n=1}^{\infty} h_{n,i}\, q^n = q^{\frac{1}{8}} \frac{\eta(\tau)^3}{\vartheta_1(\tau,z)^2} \Big(2 \sum_g c(g) \overline{\mathrm{Tr}_{R_i}(g)}\, \phi_g(\tau,z) - (\delta_{i,1} + \delta_{i,2})\, \mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=0}(\tau,z) \Big) \ , \tag{4.5}$$

where $i = 1, \ldots, 26$. In deriving (4.5) we have used that $H_{00} = R_1 \oplus R_2 = \mathbf{1} \oplus \mathbf{23}$ and $H_0 = 2 \cdot R_1 = 2 \cdot \mathbf{1}$. For $i > 2$, (4.5) simplifies, and we obtain

$$\sum_{n=1}^{\infty} h_{n,i}\, q^n = -2q^{\frac{1}{8}} \frac{\sum_g c(g) \overline{\mathrm{Tr}_{R_i}(g)}\, F_g(\tau)}{\eta(\tau)^3} = -2 \frac{\sum_g c(g) \overline{\mathrm{Tr}_{R_i}(g)}\, F_g(\tau)}{\prod_{n=1}^{\infty}(1-q^n)^3}\,, \qquad i > 2\,, \quad (4.6)$$

where we used (3.21), as well as (3.22) and (4.3) again. Furthermore, we have plugged in the explicit expression for $\phi_{-2,1}$ from (A.7).

The character values $\mathrm{Tr}_{R_i}(g)$ are given in table 3, and the explicit values of $c(g)$ are tabulated in table 2. The analysis of the previous sections provide closed formulae for the twining characters $\phi_g(\tau, z)$ for all conjugacy classes g in $\mathbb{M}_{24}$, so that the right hand side of (4.5) can be easily evaluated. Thus we can determine the multiplicities $h_{n,i}$ explicitly. The statement that there are underlying $\mathbb{M}_{24}$ representations H_n is now simply equivalent to the property of the multiplicities $h_{n,i}$ to be non-negative integers. We have worked out these multiplicities for $n \leq 500$, and all of them are indeed non-negative integers; the explicit values for $n \leq 30$ are listed in table 4.

Since all the characters we have constructed have real coefficients, the multiplicities of conjugate representations are always equal, so that all H_n are real representations, as expected. It is also remarkable that the multiplicities of the real irreducible representations in (4.1) are always even, at least, up to $n = 500$ (see also [7]). This suggests that the actual symmetry group may be slightly bigger than the Mathieu group $\mathbb{M}_{24}$.

5. Conclusions

In this paper we have accumulated compelling evidence for the conjecture of Eguchi, Ooguri and Tachikawa [1] that the states contributing to the elliptic genus of K3 carry an action of the Mathieu group $\mathbb{M}_{24}$. More specifically, we have found closed form expressions for the twining genera of K3 for all conjugacy classes of $\mathbb{M}_{24}$, thus completing the programme initiated in [7, 8]. We have shown that the twining genera transform indeed as Jacobi forms of index one and weight zero under $\Gamma_0(N)$, where N is the order of the corresponding group element. The twining genera of the conjugacy classes that have no representative contained in $\mathbb{M}_{23} \subset \mathbb{M}_{24}$ have a non-trivial multiplier system, which we have identified, see eq. (3.8).

The explicit knowledge of all twining genera allows one to determine the decomposition of the elliptic genus of K3 in terms of $\mathbb{M}_{24}$ representations, and we have checked that the multiplicities with which these representations appear are indeed non-negative integers, at least for the first 500 coefficients — see also table 4 for explicit results for $n \leq 30$. This is a highly non-trivial consistency check; indeed, if we were able to show that *all* of these multiplicities are non-negative integers, this would effectively prove the conjecture of [1].

Another, more conceptual, proof of the conjecture would consist of constructing explicitly the action of $\mathbb{M}_{24}$ on the BPS states contributing to the elliptic genus of K3. For some generators in $\mathbb{M}_{23} \subset \mathbb{M}_{24}$ this can be done fairly directly since they describe geometric automorphisms of the K3 surface [15, 16], see also [18] for some recent progress in this direction. However, it seems unlikely that these purely geometrical symmetries will suffice to account for the full $\mathbb{M}_{24}$ symmetry of the elliptic genus. In fact, one may guess that one will need to consider the full moduli space of the non-linear sigma model in order to achieve this. Thus at least some of the required symmetries may have an interpretation in terms of truly stringy symmetries, such as *e.g.* T-duality. It would be very interesting to explore these ideas in more detail [19].

On a more technical note, it is amusing to note (see also [8]) that, in contrast to the McKay-Thompson series in Monstrous Moonshine [5], not all twining genera seem to be Hauptmoduls for genus zero congruence subgroups (although many are). For example, the twining character 21AB, for which the relevant modular group, $\Gamma_0(21)$, is not genus 0. This fact is also reflected in equation (3.15) where we have given the explicit expression of $\Phi^{(3)}_{21\mathrm{AB}}(\tau)$ as a modular form of $\Gamma_0(21)$. Unlike the remaining twining characters, it cannot be written as a rational function of a single modular function (the Hauptmodul) but rather involves two functions (the McKay-Thompson series $T_{[21\mathrm{B}]}$ and $T_{[21\mathrm{D}]}$). It would be very interesting to understand what the significance of the genus zero property in this context is.

Acknowledgments

We thank Terry Gannon for useful communications and John McKay for inspiring discussions. The research of MRG and SH is partially supported by a grant from the Swiss National Science Foundation, and the research of RV is supported by an INFN Fellowship.

Class	$c(g)^{-1}$	Class	$c(g)^{-1}$	Class	$c(g)^{-1}$
1A	244823040	6A	24	4C	96
2A	21504	11A	11	3B	504
3A	1080	15AB	15	2B	7680
5A	60	14AB	14	10A	20
4B	128	23AB	23	21AB	21
7AB	42	12B	12	4A	384
8A	16	6B	24	12A	12

Table 2: Order $c(g)^{-1}$ of the centraliser of the class g.

Table 3: The character table of the Mathieu group $\mathbb{M}_{24}$. The rows correspond to the representations R_i, numbered from 1 to 26, while the columns describe the different conjugacy classes. Finally, $e_p^{\pm} = (-1 \pm i\sqrt{p})/2$.

i	1A	2A	3A	5A	4B	7A	7B	8A	6A	11A	15A	15B	14A	14B	23A	23B	12B	6B	4C	3B	2B	10A	21A	21B	4A	12A
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	23	7	5	3	3	2	2	1	1	1	0	0	0	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	252	28	9	2	4	0	0	0	1	-1	-1	-1	0	0	-1	-1	0	0	0	0	12	2	0	0	4	1
4	253	13	10	3	1	1	1	-1	-2	0	0	0	-1	-1	0	0	1	1	1	1	-11	-1	1	1	-3	0
5	1771	-21	16	1	-5	0	0	-1	0	0	1	1	0	0	0	0	-1	-1	-1	7	11	1	0	0	3	0
6	3520	64	10	0	0	-1	-1	0	-2	0	0	0	1	1	1	1	0	0	0	-8	0	0	-1	-1	0	0
7	45	-3	0	0	1	e_7^+	e_7^-	-1	0	1	0	0	$-e_7^+$	$-e_7^-$	-1	-1	1	-1	1	3	5	0	e_7^-	e_7^+	-3	0
8	45	-3	0	0	1	e_7^-	e_7^+	-1	0	1	0	0	$-e_7^-$	$-e_7^+$	-1	-1	1	-1	1	3	5	0	e_7^+	e_7^-	-3	0
9	990	-18	0	0	2	e_7^+	e_7^-	0	0	0	0	0	e_7^+	e_7^-	1	1	1	-1	-2	3	-10	0	e_7^-	e_7^+	6	0
10	990	-18	0	0	2	e_7^-	e_7^+	0	0	0	0	0	e_7^-	e_7^+	1	1	1	-1	-2	3	-10	0	e_7^+	e_7^-	6	0
11	1035	-21	0	0	3	$2e_7^+$	$2e_7^-$	-1	0	1	0	0	0	0	0	0	-1	1	-1	-3	-5	0	$-e_7^-$	$-e_7^+$	3	0
12	1035	-21	0	0	3	$2e_7^-$	$2e_7^+$	-1	0	1	0	0	0	0	0	0	-1	1	-1	-3	-5	0	$-e_7^+$	$-e_7^-$	3	0
13	1035	27	0	0	-1	-1	-1	1	0	1	0	0	-1	-1	0	0	0	2	3	6	35	0	-1	-1	3	0
14	231	7	-3	1	-1	0	0	-1	1	0	e_{15}^+	e_{15}^-	0	0	1	1	0	0	3	0	-9	1	0	0	-1	-1
15	231	7	-3	1	-1	0	0	-1	1	0	e_{15}^-	e_{15}^+	0	0	1	1	0	0	3	0	-9	1	0	0	-1	-1
16	770	-14	5	0	-2	0	0	0	1	0	0	0	0	0	e_{23}^+	e_{23}^-	1	1	-2	-7	10	0	0	0	2	-1
17	770	-14	5	0	-2	0	0	0	1	0	0	0	0	0	e_{23}^-	e_{23}^+	1	1	-2	-7	10	0	0	0	2	-1
18	483	35	6	-2	3	0	0	-1	2	-1	1	1	0	0	0	0	0	0	3	0	3	-2	0	0	3	0
19	1265	49	5	0	1	-2	-2	1	1	0	0	0	0	0	0	0	0	0	-3	8	-15	0	1	1	-7	-1
20	2024	8	-1	-1	0	1	1	0	-1	0	-1	-1	1	1	0	0	0	0	0	8	24	-1	1	1	8	-1
21	2277	21	0	-3	1	2	2	-1	0	0	0	0	0	0	0	0	0	2	-3	6	-19	1	-1	-1	-3	0
22	3312	48	0	-3	0	1	1	0	0	1	0	0	-1	-1	0	0	0	-2	0	-6	16	1	1	1	0	0
23	5313	49	-15	3	-3	0	0	-1	1	0	0	0	0	0	0	0	0	0	-3	0	9	-1	0	0	1	1
24	5796	-28	-9	1	4	0	0	0	-1	-1	1	1	0	0	0	0	0	0	0	0	36	1	0	0	-4	-1
25	5544	-56	9	-1	0	0	0	0	1	0	-1	-1	0	0	1	1	0	0	0	0	24	-1	0	0	-8	1
26	10395	-21	0	0	-1	0	0	1	0	0	0	0	0	0	-1	-1	0	0	3	0	-45	0	0	0	3	0

i	1	2	3	4	5	6	$7,8$	$9,10$	$11,12$	13	$14,15$	$16,17$	18	19	20	21	22	23	24	25	26
n	**1**	**23**	**252**	**253**	**1771**	**3520**	**45**	**990**	**1035**	**1035**	**231**	**770**	**483**	**1265**	**2024**	**2277**	**3312**	**5313**	**5796**	**5544**	**10395**
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
6	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
7	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	2	2	2	2
8	0	0	0	0	0	2	0	1	1	0	0	0	0	2	0	2	2	4	2	2	6
9	0	0	0	0	2	4	0	0	2	2	0	2	2	0	2	2	4	4	8	8	10
10	0	0	0	2	4	8	0	2	2	2	2	0	2	4	4	6	6	12	10	10	24
11	0	0	0	0	8	12	0	4	4	6	0	4	0	2	10	8	14	22	26	24	40
12	0	2	2	4	12	30	0	8	8	4	2	6	4	12	12	18	26	40	40	38	80
13	0	0	4	2	26	44	2	14	14	18	2	10	6	16	30	28	44	70	84	80	136
14	0	0	4	6	38	86	0	24	24	22	8	16	14	34	46	58	80	128	132	126	254
15	0	0	12	8	78	144	2	40	44	46	8	38	18	46	86	88	138	218	246	238	424
16	0	2	18	22	122	252	2	72	72	68	18	50	36	100	140	170	232	378	400	382	742
17	0	2	30	26	212	410	8	116	124	130	25	94	54	140	246	262	392	630	704	670	1222
18	0	6	50	58	342	704	6	194	202	192	50	148	100	256	388	454	654	1044	1120	1074	2058
19	0	4	80	72	582	1116	18	318	332	346	68	252	150	394	664	722	1062	1702	1880	1800	3320
20	0	14	128	138	904	1836	20	516	536	520	126	390	254	676	1036	1196	1716	2764	2980	2846	5408
21	2	20	214	200	1476	2902	40	814	860	872	182	652	396	1020	1684	1862	2742	4384	4828	4622	8572
22	2	32	328	346	2302	4616	55	1298	1348	1336	314	988	640	1686	2630	3000	4324	6950	7532	7204	13620
23	2	40	512	496	3638	7166	98	2020	2118	2144	460	1590	972	2546	4162	4624	6768	10856	11898	11376	21204
24	0	80	798	824	5584	11192	132	3140	3278	3236	744	2426	1544	4050	6376	7248	10500	16834	18294	17504	32976
25	8	108	1232	1208	8654	17084	234	4814	5038	5084	1106	3764	2336	6108	9892	11042	16112	25840	28288	27056	50524
26	6	174	1860	1904	13090	26148	322	7348	7670	7626	1742	5677	3602	9444	14968	16940	24566	39428	42894	41022	77176
27	12	252	2836	2802	19914	39436	514	11092	11618	11666	2560	8688	5394	14100	22744	25462	37148	59564	65114	62294	116494
28	16	398	4238	4310	29772	59330	742	16686	17418	17356	3922	12912	8160	21414	34026	38434	55764	89490	97456	93218	175146
29	26	560	6328	6286	44512	88280	1154	24840	25994	26078	5758	19380	12090	31636	50892	57068	83146	133356	145690	139342	260828
30	34	876	9368	9486	65776	131020	1642	36824	38480	38368	8642	28580	18008	47172	75158	84776	123176	197596	215318	205970	386724

Table 4: Multiplicities $h_{n,i}$ in the decomposition of the representations $H_n = \oplus_{i=1}^{26} h_{n,i} R_i$, for $n \leq 30$. The representations R_i are numbered as in table 3, in the second row the dimension is given. Pairs of conjugate representations are listed together, because they always appear with the same multiplicities.

A. Definitions

Our conventions for the Dedekind eta and the Jacobi theta functions are

$$\begin{aligned}
\eta(\tau) &= q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n) \\
\vartheta_1(\tau,z) &= -iq^{\frac{1}{8}} y^{\frac{1}{2}} \prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^{n-1}) \\
\vartheta_2(\tau,z) &= 2\,q^{\frac{1}{8}} \cos(\pi z) \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^n)(1+y^{-1}q^n) \\
\vartheta_3(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^{n-1/2})(1+y^{-1}q^{n-1/2}) \\
\vartheta_4(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1-yq^{n-1/2})(1-y^{-1}q^{n-1/2}) \quad .
\end{aligned} \tag{A.1}$$

Under modular transformations the ϑ and η functions transform as

$$\vartheta_1(\tau+1,z) = e^{\frac{2\pi i}{8}}\,\vartheta_1(\tau,z) \qquad \vartheta_1(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = -(-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_1(\tau,z)\ , \tag{A.2}$$

$$\vartheta_2(\tau+1,z) = e^{\frac{2\pi i}{8}}\,\vartheta_2(\tau,z) \qquad \vartheta_2(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_4(\tau,z)\ , \tag{A.3}$$

$$\vartheta_3(\tau+1,z) = \vartheta_4(\tau,z) \qquad \vartheta_3(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_3(\tau,z)\ , \tag{A.4}$$

$$\vartheta_4(\tau+1,z) = \vartheta_3(\tau,z) \qquad \vartheta_4(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_2(\tau,z)\ , \tag{A.5}$$

as well as

$$\eta(\tau+1) = e^{\frac{2\pi i}{24}}\,\eta(\tau) \qquad \eta(-\tfrac{1}{\tau}) = (-i\tau)^{\frac{1}{2}}\,\eta(\tau)\ . \tag{A.6}$$

The theta constants $\vartheta_a(\tau)$ are defined as $\vartheta_a(\tau) \equiv \vartheta_a(\tau, z=0)$. The standard weak Jacobi forms $\phi_{0,1}$ and $\phi_{-2,1}$ of index 1 and weight 0 and 2 can be defined as [11]

$$\phi_{0,1}(\tau,z) = 4\sum_{i=2}^{4}\frac{\vartheta_i(\tau,z)^2}{\vartheta_i(\tau,0)^2}\ , \qquad \phi_{-2,1}(\tau,z) = -\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^6}\ . \tag{A.7}$$

For completeness we give here the first few terms of the McKay-Thompson series that appear in our analysis

$$T_{[8E]} = \frac{1}{q} + 4q + 2q^3 - 8q^5 - q^7 + 20q^9 - 2q^{11} - 40q^{13} + \cdots\ , \tag{A.8}$$

$$T_{[10E]} = \frac{1}{q} + q + 2q^2 + 2q^3 - 2q^4 - q^5 - 4q^7 - 2q^8 + 5q^9 + 2q^{10} + 8q^{12} + \cdots\ , \tag{A.9}$$

$$T_{[12I]} = \frac{1}{q} + 2q + q^3 - 2q^7 - 2q^9 + 2q^{11} + 4q^{13} + 3q^{15} + \cdots\ , \tag{A.10}$$

$$T_{[21B]} = \frac{1}{q} - q - q^2 + q^3 + 2q^4 - q^5 + 3q^6 - q^7 - q^8 - 2q^9 + q^{11} + \cdots\ , \tag{A.11}$$

$$T_{[21D]} = \frac{1}{q} + 5q + 8q^2 + 16q^3 + 26q^4 + 44q^5 + 66q^6 + 104q^7 + \cdots\ , \tag{A.12}$$

see [17, 20] for more information about these series.

B. Modular forms for $\Gamma_0(N)$

In this section, we will describe a basis of the space $M_k(\Gamma_0(N))$ of modular forms of weight k under $\Gamma_0(N)$. Our main references for this section are [21, 22].

The space $M_k(\Gamma_0(N))$ of modular forms of weight k splits into a direct sum

$$M_k(\Gamma_0(N)) = E_k(\Gamma_0(N)) \oplus S_k(\Gamma_0(N)) \ , \tag{B.1}$$

where $S_k(\Gamma_0(N))$ is the space of cusp forms, *i.e.* forms they vanish at all the cusps[9] of $\overline{\mathbb{H}/\Gamma_0(N)}$. The space $E_k(\Gamma_0(N))$ is defined as the unique subspace satisfying (B.1) that is invariant under the action of all Hecke operators [22].

A convenient basis for $E_k(\Gamma_0(N))$ is given by (generalised) Eisenstein series [23]. Let

$$E_k(\tau) = -\frac{B_k}{2k} + \sum_{n=1}^{\infty}\Big(\sum_{d|n} d^{k-1}\Big)q^n \tag{B.2}$$

be the standard Eisenstein series of weight k, where B_k are the Bernoulli numbers. For $k > 2$, k even, E_k is a modular form of weight k under $SL(2,\mathbb{Z})$, whereas

$$E_2\Big(\frac{a\tau+b}{c\tau+d}\Big) = (c\tau+d)^2 E_2(\tau) - \frac{1}{4\pi i}c(c\tau+d) \ . \tag{B.3}$$

The definition of the Eisenstein series can be generalised to include modular forms under $\Gamma_0(N)$. In particular,[10]

$$\psi^{(N)} = q\frac{\partial}{\partial q}\log\frac{\eta(N\tau)}{\eta(\tau)} = E_2(\tau) - NE_2(N\tau) \tag{B.4}$$

is a modular form of weight 2 under $\Gamma_0(N)$. The (generalised) Eisenstein series

$$E_k^{\chi_m}(\tau) = \sum_{n=1}^{\infty}\Big(\sum_{d|n}\overline{\chi_m(d)}\,\chi_m(n/d)\,d^{k-1}\Big)\,q^n \ , \tag{B.5}$$

where k is even and χ_m is a non-trivial Dirichlet character of modulus m, is a modular form of weight k under $\Gamma_0(m^2)$. The only cases we need are

$$E_k^{(9)} = E_k^{\chi_3} \in M_k(\Gamma_0(9)) \ , \quad E_k^{(16)} = E_k^{\chi_4} \in M_k(\Gamma_0(16)) \ , \quad E_k^{(144)} = E_k^{\chi_{12}} \in M_k(\Gamma_0(144)) \ , \tag{B.6}$$

where χ_3, χ_4 and χ_{12} are the primitive Dirichlet characters of modulus 3, 4 and 12, that are uniquely determined by

$$\chi_3(2) = -1 \ , \qquad \chi_4(3) = -1 \ , \qquad \chi_{12}(5) = \chi_{12}(7) = -1 \ . \tag{B.7}$$

---

[9] The cusps correspond to $\Gamma_0(N)$-orbits in $\mathbb{Q}\cup\{\infty\}$, with $\Gamma_0(N)$ acting by fractional linear transformations.

[10] The Eisenstein series $\psi^{(N)}$ are related to the modular forms $\phi^{(N)}$ in [7] by $\psi^{(N)} = \frac{N-1}{24}\phi^{(N)}(\tau)$.

In general, the space $S_k(\Gamma_0(N))$ is not generated by Eisenstein series. It is obvious that if $M|N$, then $M_k(\Gamma_0(M)) \subset M_k(\Gamma_0(N))$. More generally, it is easy to see that if $f \in M_k(\Gamma_0(M))$, then, for any divisor n of N/M, $f(n\tau) \in \Gamma_0(N)$. For all $M|N$ and $n|(N/M)$, we define the map

$$\alpha_n : M_k(\Gamma_0(M)) \to M_k(\Gamma_0(N)) \tag{B.8}$$

$$f(\tau) \mapsto \alpha_n(f)(\tau) = f(n\tau) \ . \tag{B.9}$$

The map α_n sends cusp forms to cusp forms and the union of the images $\alpha_n(S_k(\Gamma_0(N/n))) \subseteq S_k(\Gamma_0(N))$, for all $n|N$, $n > 1$, is called the *old subspace* of cusp forms. The complement of the old subspace which is invariant under all Hecke operators is called the *new subspace*. Thus, we have a decomposition

$$S_k(\Gamma_0(N)) = \bigoplus_{M|N} \bigoplus_{n|(N/M)} \alpha_n(S_k(\Gamma_0(M))_{new}) \ , \tag{B.10}$$

where $S_k(\Gamma_0(M))_{new}$ is the new subspace for $\Gamma_0(M)$. A basis for $S_k(\Gamma_0(M))_{new}$ for the cases of interest are listed below; a more extended list of their coefficients can be found at `http://modi.countnumber.de/`. Some Fourier expansions have been computed using SAGE (`http://www.sagemath.org/`).

- Cusp forms $f_M \in S_2(\Gamma_0(M))_{new}$ — if there is more than one generator, the different generators are denoted by $f_{M,a}, f_{M,b}, \ldots$.

$$f_{21}(\tau) = q - q^2 + q^3 - q^4 - 2q^5 - q^6 - q^7 + 3q^8 + q^9 + 2q^{10} + \cdots \tag{B.11}$$

$$f_{23,a}(\tau) = q - q^3 - q^4 - 2q^6 + 2q^7 - q^8 + 2q^9 + 2q^{10} + \cdots \tag{B.12}$$

$$f_{23,b}(\tau) = -q^2 + 2q^3 + q^4 - 2q^5 - q^6 - 2q^7 + 2q^8 + 2q^{10} + \cdots \tag{B.13}$$

$$f_{24}(\tau) = q - q^3 - 2q^5 + q^9 + 4q^{11} - 2q^{13} + 2q^{15} + 2q^{17} + \cdots \tag{B.14}$$

$$f_{48}(\tau) = q + q^3 - 2q^5 + q^9 - 4q^{11} - 2q^{13} - 2q^{15} + 2q^{17} + \cdots \tag{B.15}$$

$$f_{63,a}(\tau) = q + q^2 - q^4 + 2q^5 - q^7 - 3q^8 + 2q^{10} - 4q^{11} - 2q^{13} + \cdots \tag{B.16}$$

$$f_{63,b}(\tau) = q + q^4 + q^7 - 6q^{10} + 2q^{13} - 5q^{16} + \cdots \tag{B.17}$$

$$f_{63,c}(\tau) = q^2 - 2q^5 - q^8 + 2q^{11} + q^{14} + 2q^{17} + \cdots \tag{B.18}$$

$$f_{72}(\tau) = q + 2q^5 - 4q^{11} - 2q^{13} - 2q^{17} - 4q^{19} + \cdots \tag{B.19}$$

$$f_{144,a}(\tau) = q + 4q^7 + 2q^{13} - 8q^{19} - 5q^{25} + 4q^{31} + \cdots \tag{B.20}$$

$$f_{144,b}(\tau) = q + 2q^5 + 4q^{11} - 2q^{13} - 2q^{17} + 4q^{19} + \cdots \ . \tag{B.21}$$

- Cusp forms $g_M \in S_4(\Gamma_0(M))_{new}$

$$g_5(\tau) = q - 4q^2 + 2q^3 + 8q^4 - 5q^5 - 8q^6 + 6q^7 - 23q^9 + \cdots \tag{B.22}$$

$$g_6(\tau) = q - 2q^2 - 3q^3 + 4q^4 + 6q^5 + 6q^6 - 16q^7 - 8q^8 + 9q^9 + \cdots \tag{B.23}$$

$$g_8(\tau) = q - 4q^3 - 2q^5 + 24q^7 - 11q^9 - 44q^{11} + 22q^{13} + \cdots \tag{B.24}$$

$$g_{10}(\tau) = q + 2q^2 - 8q^3 + 4q^4 + 5q^5 - 16q^6 - 4q^7 + 8q^8 + 37q^9 + \cdots \tag{B.25}$$

$$g_{12}(\tau) = q + 3q^3 - 18q^5 + 8q^7 + 9q^9 + 36q^{11} - 10q^{13} - 54q^{15} + \cdots \ . \tag{B.26}$$

- Cusp forms $h_M \in S_6(\Gamma_0(M))_{new}$

$$\begin{aligned}
h_3(\tau) =& q - 6q^2 + 9q^3 + 4q^4 + 6q^5 - 54q^6 - 40q^7 + 168q^8 + \cdots && \text{(B.27)}\\
h_{7,a}(\tau) =& q - 10q^2 - 14q^3 + 68q^4 - 56q^5 + 140q^6 - 49q^7 - 360q^8 + \cdots && \text{(B.28)}\\
h_{7,b}(\tau) =& q + 4q^2 - 2q^4 - 14q^5 - 84q^6 + 49q^7 - 10q^8 + \cdots && \text{(B.29)}\\
h_{7,c}(\tau) =& q^2 - 6q^3 + 9q^4 + 10q^5 - 30q^6 + 11q^8 + \cdots && \text{(B.30)}\\
h_{21,a}(\tau) =& q + q^2 - 9q^3 - 31q^4 - 34q^5 - 9q^6 - 49q^7 - 63q^8 + \cdots && \text{(B.31)}\\
h_{21,b}(\tau) =& q + 5q^2 + 9q^3 - 7q^4 + 94q^5 + 45q^6 - 49q^7 - 195q^8 + \cdots && \text{(B.32)}\\
h_{21,c}(\tau) =& q - 6q^2 - 9q^3 + 4q^4 + 78q^5 + 54q^6 + 49q^7 + 168q^8 + \cdots && \text{(B.33)}\\
h_{21,d}(\tau) =& q + 10q^2 + 9q^3 + 68q^4 - 106q^5 + 90q^6 - 49q^7 + 360q^8 + \cdots \ . && \text{(B.34)}
\end{aligned}$$

- Cusp forms $k_M \in S_8(\Gamma_0(M))_{new}$

$$k_2(\tau) = q - 8q^2 + 12q^3 + 64q^4 - 210q^5 - 96q^6 + 1016q^7 - 512q^8 + \cdots \ . \qquad \text{(B.35)}$$

- Cusp forms $l_M \in S_{12}(\Gamma_0(M))_{new}$

$$\begin{aligned}
l_3(\tau) =& q + 78q^2 - 243q^3 + 4036q^4 - 5370q^5 - 18954q^6 - 27760q^7 + \cdots && \text{(B.36)}\\
l_{6,a}(\tau) =& q + 32q^2 + 243q^3 + 1024q^4 + 3630q^5 + 7776q^6 + 32936q^7 + \cdots && \text{(B.37)}\\
l_{6,b}(\tau) =& q - 32q^2 - 243q^3 + 1024q^4 + 5766q^5 + 7776q^6 + 72464q^7 + \cdots && \text{(B.38)}\\
l_{6,c}(\tau) =& q - 32q^2 + 243q^3 + 1024q^4 - 11730q^5 - 7776q^6 - 50008q^7 + \cdots \ . && \text{(B.39)}
\end{aligned}$$

To summarise, $M_k(\Gamma_0(N))$ is thus generated by

$$\begin{array}{lll}
k=2: & \psi^{(n)}(\tau) & \text{for } n|N\,,\ n>1 \\
 & E_2^{\chi_m}(n\tau) & \text{for } nm^2|N\,,\ m>1 \\
 & f_{M,a}(n\tau),\ f_{M,b}(n\tau),\ \ldots & \text{for } nM|N \\
 & & \\
k=4: & E_4(n\tau) & \text{for } n|N \\
 & E_4^{\chi_m}(n\tau) & \text{for } nm^2|N\,,\ m>1 \\
 & g_{M,a}(n\tau),\ g_{M,b}(n\tau),\ \ldots & \text{for } nM|N\ ,
\end{array} \qquad \text{(B.40)}$$

and the cases $k = 6, 8, \ldots$ *etc.* are similar to $k = 4$, with $g_{M,a}$ replaced by $h_{M,a}$, *etc.*

C. The modular forms of weight two

With the conventions of the previous appendix we can now give the explicit expressions for the modular forms $F_g \in M_2(\Gamma_0(Nh))$, as well as for $F_g^h \in M_{2h}(\Gamma_0(N))$. For the classes in (3.2) for which $h = 1$, the relevant formulae are [7]

$$\begin{aligned}
F_{2\mathrm{A}}(\tau) &= 16\psi^{(2)}(\tau)\\
F_{3\mathrm{A}}(\tau) &= 9\psi^{(3)}(\tau)\\
F_{4\mathrm{B}}(\tau) &= -4\psi^{(2)}(\tau) + 8\psi^{(4)}(\tau)\\
F_{5\mathrm{A}}(\tau) &= 5\psi^{(5)}(\tau)\\
F_{6\mathrm{A}}(\tau) &= -2\psi^{(2)}(\tau) - 3\psi^{(3)}(\tau) + 6\psi^{(6)}(\tau)\\
F_{7\mathrm{AB}}(\tau) &= \frac{7}{2}\psi^{(7)}(\tau)\\
F_{8\mathrm{A}}(\tau) &= -2\psi^{(4)}(\tau) + 4\psi^{(8)}(\tau)\ ,\\
F_{11\mathrm{A}}(\tau) &= \frac{11}{5}\big(\psi^{(11)}(\tau) - \eta(\tau)^2\eta(11\tau)^2\big)\\
F_{14\mathrm{AB}}(\tau) &= \frac{1}{6}\big(-2\psi^{(2)}(\tau) - 7\psi^{(7)}(\tau) + 14\psi^{(14)}(\tau) - 14\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau)\big)\\
F_{15\mathrm{AB}}(\tau) &= \frac{1}{8}\big(-3\psi^{(3)}(\tau) - 5\psi^{(5)}(\tau) + 15\psi^{(15)}(\tau) - 15\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau)\big)\\
F_{23\mathrm{AB}} &= \frac{23}{22}\big(\psi^{(23)}(\tau) - f_{23,a}(\tau) + 3f_{23,b}(\tau)\big)\ .
\end{aligned} \tag{C.1}$$

In the remaining cases we have found — the formulae for $F_{2\mathrm{B}}$ and $F_{4\mathrm{A}}$ were already obtained in [7]

- $g = 2\mathrm{B}$, $h = 2$

$$F_{2\mathrm{B}}(\tau) = -24\psi^{(2)}(\tau) + 16\psi^{(4)}(\tau) \qquad \in M_2(\Gamma_0(4)) \tag{C.2}$$

$$F_{2\mathrm{B}}(\tau)^2 = -16E_4(\tau) + 256E_4(2\tau) \qquad \in M_4(\Gamma_0(2)) \tag{C.3}$$

- $g = 3\mathrm{B}$, $h = 3$

$$F_{3B}(\tau) = -6\psi^{(3)}(\tau) + \frac{9}{2}\psi^{(9)}(\tau) - \frac{9}{2}E_2^{(9)}(\tau) \qquad \in M_2(\Gamma_0(9)) \tag{C.4}$$

$$F_{3B}(\tau)^3 = -1944\psi^{(3)}(\tau)^3 + \frac{45}{2}E_6(\tau) - \frac{2187}{2}E_6(3\tau) \qquad \in M_6(\Gamma_0(3)) \tag{C.5}$$

- $g = 4\mathrm{A}$, $h = 2$

$$F_{4\mathrm{A}}(\tau) = 4\psi^{(2)}(\tau) - 12\psi^{(4)}(\tau) + 8\psi^{(8)}(\tau) \qquad \in M_2(\Gamma_0(8)) \tag{C.6}$$

$$F_{4\mathrm{A}}(\tau)^2 = -16E_4(2\tau) + 256E_4(4\tau) \qquad \in M_4(\Gamma_0(4)) \tag{C.7}$$

- $g = 4\mathrm{C}$, $h = 4$

$$F_{4\mathrm{C}}(\tau) = 2\psi^{(4)}(\tau) - 6\psi^{(8)}(\tau) + 4\psi^{(16)}(\tau) - 4E_2^{(16)}(\tau) \qquad \in M_2(\Gamma_0(16)) \tag{C.8}$$

$$F_{4\mathrm{C}}(\tau)^4 = -\tfrac{32}{17}E_8(2\tau) + \tfrac{8192}{17}E_8(4\tau) - 16k_2(\tau) - \tfrac{512}{17}k_2(2\tau) \qquad \in M_8(\Gamma_0(4)) \tag{C.9}$$

- $g = 6\mathrm{B}$, $h = 6$

$$\begin{aligned} F_{6\mathrm{B}}(\tau) = &-\tfrac{3}{2}\psi^{(2)}(\tau) - 2\psi^{(3)}(\tau) + \psi^{(4)}(\tau) + 6\psi^{(6)}(\tau) + \tfrac{3}{2}\psi^{(9)}(\tau) - 4\psi^{(12)}(\tau) \\ &- \tfrac{9}{2}\psi^{(18)}(\tau) + 3\psi^{(36)}(\tau) - \tfrac{3}{2}E_2^{(9)}(\tau) \\ &- 9E_2^{(9)}(2\tau) - 12E_2^{(9)}(4\tau) \qquad \in M_2(\Gamma_0(36)) \end{aligned} \tag{C.10}$$

$$\begin{aligned} F_{6\mathrm{B}}(\tau)^6 = &\, c_1 E_{12}(\tau) + c_2 E_{12}(2\tau) + c_3 E_{12}(3\tau) + c_4 E_{12}(6\tau) + c_5 \Delta(\tau) \\ &+ c_6 \Delta(2\tau) + c_7 \Delta(3\tau) + c_8 \Delta(6\tau) + c_9 l_{6,a}(\tau) \\ &+ c_{10} l_{6,b}(\tau) + c_{11} l_{6,c}(\tau) + c_{12} l_3(\tau) + c_{13} l_3(2\tau) \qquad \in M_{12}(\Gamma_0(6)) \end{aligned} \tag{C.11}$$

where $\Delta(\tau) = ((240E_4(\tau))^3 - (504E_6(\tau)^2))/1728$ and

p	1	2	3	4	5	6	7	8	9	10	11	12	13
c_p	$\frac{1}{22951565}$	$-\frac{4096}{22951565}$	$-\frac{531441}{22951565}$	$\frac{2176782336}{22951565}$	$-\frac{189277}{58735}$	$-\frac{91776}{3455}$	$\frac{37712628}{58735}$	$\frac{27713664}{3455}$	$-\frac{297}{140}$	$-\frac{308}{145}$	$-\frac{95}{52}$	$-\frac{194697}{71978}$	$\frac{304128}{2117}$

- $g =$10A, $h = 2$

$$\begin{aligned} F_{10\mathrm{A}}(\tau) = &\, \psi^{(2)}(\tau) - \frac{2}{3}\psi^{(4)}(\tau) + \frac{5}{3}\psi^{(5)}(\tau) - 5\psi^{(10)}(\tau) \\ &+ \frac{10}{3}\psi^{(20)}(\tau) - \frac{10}{3}\eta(2\tau)^2\eta(10\tau)^2 \qquad \in M_2(\Gamma_0(20)) \end{aligned} \tag{C.12}$$

$$\begin{aligned} F_{10\mathrm{A}}(\tau)^2 = &\, \frac{1}{39}E_4(\tau) - \frac{16}{39}E_4(2\tau) - \frac{625}{39}E_4(5\tau) + \frac{10000}{39}E_4(10\tau) \\ &- \frac{10}{3}g_{10}(\tau) - \frac{35}{13}g_5(\tau) + \frac{40}{13}g_5(2\tau) \qquad \in M_4(\Gamma_0(10)) \end{aligned} \tag{C.13}$$

- $g = 12\mathrm{A}$, $h = 2$

$$\begin{aligned} F_{12\mathrm{A}}(\tau) = &-\frac{1}{2}\psi^{(2)}(\tau) + \frac{3}{2}\psi^{(4)}(\tau) + \frac{3}{2}\psi^{(6)}(\tau) - \psi^{(8)}(\tau) \\ &- \frac{9}{2}\psi^{(12)}(\tau) + 3\psi^{(24)}(\tau) - 3f_{24}(\tau) \qquad \in M_2(\Gamma_0(24)) \end{aligned} \tag{C.14}$$

$$\begin{aligned} F_{12\mathrm{A}}(\tau)^2 = &\, \frac{1}{5}E_4(2\tau) - \frac{16}{5}E_4(4\tau) - \frac{81}{5}E_4(6\tau) + \frac{1296}{5}E_4(12\tau) \\ &- 3g_{12}(\tau) - 3g_6(\tau) + \frac{24}{5}g_6(2\tau) \qquad \in M_4(\Gamma_0(12)) \end{aligned} \tag{C.15}$$

- $g = 12\mathrm{B}$, $h = 12$

$$\begin{aligned} F_{12\mathrm{B}}(\tau) = {}& \frac{1}{8}\psi^{(4)}(\tau) - \frac{3}{8}\psi^{(8)}(\tau) - \frac{1}{2}\psi^{(12)}(\tau) + \frac{1}{4}\psi^{(16)}(\tau) + \frac{3}{2}\psi^{(24)}(\tau) + \frac{3}{8}\psi^{(36)}(\tau) \\ & - \psi^{(48)}(\tau) - \frac{9}{8}\psi^{(72)}(\tau) + \frac{3}{4}\psi^{(144)}(\tau) - \frac{3}{2}E_2^{(9)}(4\tau) - 9E_2^{(9)}(8\tau) - 12E_2^{(9)}(16\tau) \\ & - \frac{1}{4}E_2^{(16)}(\tau) - 3E_2^{(16)}(3\tau) - \frac{27}{4}E_2^{(16)}(9\tau) - \frac{3}{4}E_2^{(144)}(\tau) - \frac{3}{2}f_{24}(2\tau) \\ & - \frac{27}{2}f_{24}(6\tau) - \frac{3}{4}f_{48}(\tau) + \frac{27}{4}f_{48}(3\tau) + \frac{9}{2}f_{72}(2\tau) - \frac{9}{4}f_{144,b}(\tau) \ \in M_2(\Gamma_0(144)) \end{aligned} \tag{C.16}$$

We have also worked out the expansion of $F^{12}_{12\mathrm{B}}$ in terms of generators of $M_{24}(\Gamma_0(12))$, but since the final expression is exceedingly complicated, we shall not spell it out here.

- $g = 21\mathrm{AB}$, $h = 3$

$$\begin{aligned} F_{21\mathrm{AB}}(\tau) = {}& \frac{1}{8}\psi^{(3)}(\tau) + \frac{7}{32}\psi^{(7)}(\tau) - \frac{3}{32}\psi^{(9)}(\tau) - \frac{7}{8}\psi^{(21)}(\tau) + \frac{21}{32}\psi^{(63)}(\tau) + \frac{3}{32}E_2^{(9)}(\tau) \\ & - \frac{147}{32}E_2^{(9)}(7\tau) - \frac{21}{32}f_{21}(\tau) + \frac{189}{32}f_{21}(3\tau) - \frac{63}{32}f_{63,a}(\tau) \qquad \in M_2(\Gamma_0(63)) \end{aligned} \tag{C.17}$$

$$\begin{aligned} F_{21\mathrm{AB}}(\tau)^3 = {}& -\tfrac{1}{169936}\,E_6(\tau) + \tfrac{729}{169936}\,E_6(3\tau) + \tfrac{117649}{169936}\,E_6(7\tau) - \tfrac{85766121}{169936}E_6\,(21\tau) \\ & - \tfrac{85641}{36608}\,h_3(\tau) - \tfrac{352947}{36608}\,h_3(7\tau) - \tfrac{64491}{94600}\,h_{7a}(\tau) - \tfrac{11263}{9728}\,h_{7b}(\tau) - \tfrac{18277}{4864}\,h_{7c}(\tau) \\ & - \tfrac{2740311}{189200}\,h_{7a}(3\tau) + \tfrac{66339}{9728}\,h_{7b}(3\tau) - \tfrac{250047}{4864}\,h_{7c}(3\tau) - \tfrac{441}{352}\,h_{21a}(\tau) \\ & - \tfrac{1323}{8800}\,h_{21b}(\tau) - \tfrac{735}{512}\,h_{21c}(\tau) - \tfrac{3087}{6400}\,h_{21d}(\tau) \qquad \in M_6(\Gamma_0(21))\ . \end{aligned} \tag{C.18}$$

References

[1] T. Eguchi, H. Ooguri and Y. Tachikawa, *Notes on the K3 surface and the Mathieu group M_{24}*, Exper. Math. **20**, 91 (2011) `[arXiv:1004.0956 [hep-th]]`.

[2] T. Eguchi and K. Hikami, *Superconformal algebras and mock theta functions 2. Rademacher expansion for K3 surface*, Commun. Number Theory Phys. **3**, 531 (2009) `[arXiv:0904.0911 [math-ph]]`.

[3] I.B. Frenkel, J. Lepowsky and A. Meurman, *Vertex Operator Algebras and the Monster*, Pure and Applied Math. **134**, Academic Press (1988).

[4] J.G. Thompson, *Some numerology between the Fischer-Griess Monster and the elliptic modular function*, Bull. Lond. Math. Soc. **11**, 352 (1979).

[5] J.H. Conway and S. Norton, *Monstrous moonshine*, Bull. Lond. Math. Soc. **11**, 308 (1979).

[6] T. Gannon, *Moonshine beyond the Monster: The Bridge connecting Algebra, Modular Forms and Physics*, Cambridge University Press (2006).

[7] M.C.N. Cheng, *K3 Surfaces, N=4 dyons, and the Mathieu group M_{24}*, Commun. Number Theory Phys. **4**, 623 (2010) `[arXiv:1005.5415 [hep-th]]`.

[8] M. R. Gaberdiel, S. Hohenegger, R. Volpato, *Mathieu twining characters for K3,* JHEP **1009**, 058 (2010) `[arXiv:1006.0221 [hep-th]]`.

[9] T. Kawai, Y. Yamada, and S.-K. Yang, *Elliptic genera and* $\mathcal{N}=2$ *superconformal field theory*, Nucl. Phys. B **414**, 191 (1994) `[arXiv:hep-th/9306096]`.

[10] J.R. David, D.P. Jatkar and A. Sen, *Product representation of dyon partition function in CHL models*, JHEP **0606**, 064 (2006) `[arXiv:hep-th/0602254]`.

[11] M. Eichler and D. Zagier, *The Theory of Jacobi Forms*, Birkhäuser (1985).

[12] T. Eguchi, H. Ooguri, A. Taormina and S.K. Yang, *Superconformal algebras and string compactification on manifolds with* $SU(N)$ *holonomy*, Nucl. Phys. B **315**, 193 (1989).

[13] T. Eguchi and A. Taormina, *Character formulas for the* $\mathcal{N}=4$ *superconformal algebra*, Phys. Lett. B **200**, 315 (1988).

[14] T. Eguchi and A. Taormina, *On the unitary representations of* $N=2$ *and* $N=4$ *superconformal algebras*, Phys. Lett. B **210**, 125 (1988).

[15] S. Mukai, *Finite groups of automorphisms of* $K3$ *surfaces and the Mathieu group*, Invent. Math. **94**, 183 (1988).

[16] S. Kondo, *Niemeier lattices, Mathieu groups and finite groups of symplectic automorphisms of K3 surfaces*, Duke Math. Journal **92**, 593 (1998), appendix by S. Mukai.

[17] D. Ford, J. McKay and S. Norton, *More on replicable functions*, Commun. Algebra **22**, 5175 (1994).

[18] A. Taormina and K. Wendland, *The symmetries of the tetrahedral Kummer surface in the Mathieu group* $\mathbb{M}_{24}$, `arXiv:1008.0954 [hep-th]`.

[19] M.R. Gaberdiel, S. Hohenegger and R. Volpato, work in progress.

[20] `http://www.research.att.com/~njas/sequences/`

[21] S. Lang, *Introduction to Modular Forms*, Grundlehren der Mathematischen Wissenschaften **222**, Springer Verlag, Berlin (1995).

[22] W. Stein, *Modular Forms, a Computational Approach*, Graduate Studies in Mathematics **79**, American Mathematical Society, Providence, RI (2007).

[23] T. Miyake, *Modular Forms*, Springer Verlag, Berlin (2006).

[24] T. Eguchi, K. Hikami, *Note on Twisted Elliptic Genus of K3 Surface,* Phys. Lett. B **694**, 446 (2011) `[arXiv:1008.4924 [hep-th]]`.

arXiv:0907.4529v3 [math.RT] 12 Apr 2012

Rademacher sums, moonshine and gravity

John F. R. Duncan*
Igor B. Frenkel†

1 March 2012

Abstract

In 1939 Rademacher derived a conditionally convergent series expression for the elliptic modular invariant, and used this expression—the first Rademacher sum—to verify its modular invariance. By generalizing Rademacher's approach we construct bases for the spaces of automorphic integrals of arbitrary even integer weight, for groups commensurable with the modular group. Our methods provide explicit expressions for the Fourier expansions of the Rademacher sums we construct at arbitrary cusps, and illuminate various aspects of the structure of the spaces of automorphic integrals, including the actions of Hecke operators.

We give a moduli interpretation for a class of groups commensurable with the modular group which includes all those that are associated to the Monster via monstrous moonshine. We show that within this class the monstrous groups can be characterized just in terms of the behavior of their Rademacher sums. In particular, the genus zero property of monstrous moonshine is encoded naturally in the properties of Rademacher sums.

Just as the ellptic modular invariant gives the graded dimension of the moonshine module, the exponential generating function of the Rademacher sums associated to the modular group furnishes the bi-graded dimension of the Verma module for the Monster Lie algebra. This result generalizes naturally to all the groups of monstrous moonshine, and recovers a certain family of monstrous Lie algebras recently introduced by Carnahan.

Our constructions suggest conjectures relating monstrous moonshine to a distinguished family of chiral three dimensional quantum gravities, and relating monstrous Lie algebras and their Verma modules to the second quantization of this family of chiral three dimensional quantum gravities.

---

*Case Western Reserve University, Department of Mathematics, 10900 Euclid Avenue, Cleveland, OH 44106.
†Yale University, Department of Mathematics, 10 Hillhouse Avenue, New Haven, CT 06520,
The research of I.F. was supported by NSF grant DMS-0457444.

Contents

1 Introduction 3
1.1 Monstrous moonshine 3
1.2 Rademacher sums 4
1.3 Conjugate Rademacher sums 6
1.4 Solid tori 7
1.5 Moonshine via Rademacher sums 8
1.6 Modified Rademacher sums 9
1.7 Hecke operators 13
1.8 Monstrous Lie algebras 15
1.9 Chiral gravity 16

2 Conventions 18
2.1 Functions 18
2.2 Divisors 19
2.3 Isometries 19
2.4 Cosets 22
2.5 Groups 23
2.6 Scalings 25
2.7 Integrals 28

3 Rademacher sums 32
3.1 Construction 32
3.2 Coefficients 36
3.3 Convergence 44
3.4 Variance 49

4 Modified Rademacher sums 56
4.1 Construction 56
4.2 Coefficients 59
4.3 Dirichlet series 60
4.4 Convergence 61
4.5 Variance 68

5 Structural applications 70
5.1 Constants 70
5.2 Inner products 71
5.3 Branching 72

5.4 Fractional orders 77
5.5 Hecke operators 81

6 Moonshine **85**
6.1 Genera 86
6.2 Moduli 87
6.3 Exponents 91
6.4 Cusps 92
6.5 Rademacher sums and the Monster 94

7 Gravity **95**
7.1 First conjecture 95
7.2 Monstrous Lie algebras 98
7.3 Second conjecture 103

References **108**

1 Introduction

1.1 Monstrous moonshine

A remarkable coincidence between the coefficients of the elliptic modular invariant

$$J(z) = \mathrm{e}(-z) + \sum_{n>0} c(n)\,\mathrm{e}(nz), \quad \mathrm{e}(nz) = \exp(2\pi \mathbf{i} nz), \tag{1.1.1}$$

and the dimensions of the irreducible representations of the *Monster group*, denoted $\mathbb{M}$, led McKay and Thompson [Tho79b] to conjecture the existence of a naturally defined infinite dimensional representation

$$V = V_{-1} \oplus V_1 \oplus V_2 \oplus \cdots \tag{1.1.2}$$

for the Monster group, with the property that $\dim V_n = c(n)$. Thompson [Tho79a] also proposed to consider the modular properties of the series

$$T_g(z) = \mathrm{e}(-z) + \sum_{n>0} (\mathrm{tr}|_{V_n} g)\,\mathrm{e}(nz), \tag{1.1.3}$$

for any $g \in \mathbb{M}$, the function $J(z)$ of (1.1.1) being recovered upon taking $g = e$ in (1.1.3). It is clear from the presentation (1.1.3) that $T_g(z)$ is invariant under the *translation group*, which we denote $B(\mathbb{Z})$ and which is generated by the translation $z \mapsto z + 1$. In the case (1.1.1) that g is the identity element, the invariance extends to the full modular group $PSL_2(\mathbb{Z})$. This leads us to the question: what is special about the invariance groups Γ_g of the McKay–Thompson series (1.1.3)?

In their remarkable paper [CN79], Conway and Norton collected an overwhelming number of coincidences and, in particular, formulated the *moonshine conjecture*:

> The series $T_g(z)$ is the normalized hauptmodul of a genus zero group Γ_g lying between $\Gamma_0(N)$ and its normalizer in $PSL_2(\mathbb{R})$, for each $g \in \mathbb{M}$.

The conjecture of McKay and Thompson was proven in [FLM88] by the explicit construction of a vertex operator algebra $V^\natural$ invariant under the Monster group. It was also shown in [FLM88] that the McKay–Thompson series $T_g(z)$ satisfy the moonshine conjecture of Conway and Norton for all elements g in a subgroup of $\mathbb{M}$ arising as an involution centralizer. Finally, a complete proof of the moonshine conjecture was given by Borcherds in [Bor92]. However, in spite of all the subsequent spectacular developments related to the Monster, the consensus of the experts (cf. e.g. [CMS04]), is that "the real nature of moonshine is still remote".

1.2 Rademacher sums

In this paper we shed new light on the properties of the McKay–Thompson series associated to the Monster, including the mysterious genus zero phenomena. It turns out that the McKay–Thompson series admit presentations as sums over the coset space $B(\mathbb{Z})\backslash\Gamma_g$. The prototype for this is the following expression for the modular invariant $J(z)$ which was originally derived by Rademacher in [Rad39].

$$J(z) + 12 = \mathrm{e}(-z) + \lim_{K\to\infty} \sum_{\substack{0<c<K \\ -K^2<d<K^2 \\ (c,d)=1}} \mathrm{e}\left(-\frac{az+b}{cz+d}\right) - \mathrm{e}\left(-\frac{a}{c}\right) \tag{1.2.1}$$

The integers a and b occurring in each summand of (1.2.1) are chosen so that $ad - bc = 1$. This expression (1.2.1) is remarkably simple, but the convergence is rather subtle. (See [Kno90] for a nice exposition.) The subtraction of the constant $\mathrm{e}(-a/c)$ in (1.2.1) ensures the existence of the limit in what is otherwise a highly divergent series. The Rademacher sum (1.2.1) has been generalized to various discrete subgroups of $PSL_2(\mathbb{R})$ in a series of papers by Knopp (cf. [Kno61a], [Kno61b], [Kno62a], [Kno62b]).

Given $\Gamma < PSL_2(\mathbb{R})$ we write Γ_∞ for the subgroup of Γ that fixes ∞. If Γ is commensurable with the modular group $PSL_2(\mathbb{Z})$ and has the property that $\Gamma_\infty = B(\mathbb{Z})$, then we may naturally associate a Rademacher sum to Γ by setting

$$R_{\Gamma}^{(-1)}(z) = \mathrm{e}(-z) + \lim_{K\to\infty} \sum_{\gamma\in(B(\mathbb{Z})\backslash\Gamma)^{\times}_{\leq K}} \mathrm{e}(-\gamma\cdot z) - \mathrm{e}(-\gamma\cdot\infty), \tag{1.2.2}$$

where the sum here is taken over the *rectangle*

$$(B(\mathbb{Z})\backslash\Gamma)^{\times}_{\leq K} = \left\{ \gamma = B(\mathbb{Z}) \begin{bmatrix} * & * \\ c & d \end{bmatrix} \in B(\mathbb{Z})\backslash\Gamma \mid 0 < c \leq K,\ -K^2 \leq d \leq K^2 \right\} \tag{1.2.3}$$

(cf. §3.1). In (1.2.3) we write $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ for the image in $PSL_2(\mathbb{R})$ of a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $SL_2(\mathbb{R})$, and in (1.2.2) we write $\gamma\cdot z$ for $(az+b)/(cz+d)$ and $\gamma\cdot\infty$ for a/c when $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$.

In general the series $R_{\Gamma}^{(-1)}(z)$ defines an abelian integral for Γ, meaning that we have

$$R_{\Gamma}^{(-1)}(\gamma\cdot z) = R_{\Gamma}^{(-1)}(z) + \omega(\gamma) \tag{1.2.4}$$

for some function $\omega : \Gamma \to \mathbb{C}$. We show in §6.1 that the Rademacher sum $R_{\Gamma}^{(-1)}(z)$ is Γ-invariant if and only if Γ has genus zero. Even more than this, we show (also in §6.1) that when Γ has genus zero the function

$$R_{\Gamma}^{(-1)}(z) - \frac{1}{2}\mathrm{c}_{\Gamma}(-1,0) \tag{1.2.5}$$

is the normalized hauptmodul for Γ (cf. §2.5), for a certain constant $\mathrm{c}_{\Gamma}(-1,0)$. This constant $\mathrm{c}_{\Gamma}(-1,0)$ turns out to be the *Rademacher constant associated to* Γ as defined in [Nor84] in the case that Γ has genus zero (cf. §5.1).

Rademacher's proof of the validity of the presentation (1.2.1) relies upon explicit formulas for the Fourier coefficients of $J(z)$. These formulas may be given in terms of Kloosterman sums and Bessel functions, and are recovered from the expression

$$\mathrm{c}_{\Gamma}(m,n) = \sum_{\gamma\in B(\mathbb{Z})\backslash\Gamma/B(\mathbb{Z})} \mathrm{Kl}_{\gamma}(m,n)\,\mathrm{Bl}_{\gamma}(m,n) \tag{1.2.6}$$

upon taking $\Gamma = PSL_2(\mathbb{Z})$ and $m = -1$, and allowing n to range over $\mathbb{Z}$. Writing $c(\gamma)$ for $|c|$ when $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, the functions $\mathrm{Kl}_{\gamma}(m,n)$ and $\mathrm{Bl}_{\gamma}(m,n)$ are defined as follows (cf. §3.2), for $\gamma \in \Gamma$.

$$\mathrm{Kl}_{\gamma}(m,n) = \mathrm{e}(m\gamma\cdot\infty)\,\mathrm{e}(-n\gamma^{-1}\cdot\infty) \tag{1.2.7}$$

$$\mathrm{Bl}_{\gamma}(m,n) = 2\pi\mathbf{i}\,\mathrm{Res}_{\xi=0}\,\mathrm{e}(mc(\gamma)^{-2}\xi^{-1})\,\mathrm{e}(n\xi)\mathrm{d}\xi \tag{1.2.8}$$

Note that $\mathrm{Kl}_{\gamma}(m,n)$ and $\mathrm{Bl}_{\gamma}(m,n)$ are well-defined complex numbers only when $c(\gamma) \neq 0$, and in this case

$$\mathrm{Kl}_\gamma(m,n) = \mathrm{e}\left(\frac{ma+nd}{c}\right), \quad \mathrm{Bl}_\gamma(m,n) = \sum_{k\geq 0}\left(\frac{2\pi\mathbf{i}}{c}\right)^{2k+2}\frac{m^{k+1}}{(k+1)!}\frac{n^k}{k!}, \tag{1.2.9}$$

for $\gamma = \left[\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right]$ and $c > 0$. The expression (1.2.6) with $\Gamma = PSL_2(\mathbb{Z})$ and $m = -1$ is, up to elementary transformations, the formula for the n-th coefficient of the elliptic modular invariant given originally by Rademacher in [Rad38]. More generally, the formula (1.2.6) recovers the Fourier coefficients (other than the constant term) of the Rademacher sum $R_\Gamma^{(-1)}(z)$, for Γ an arbitrary group commensurable with $PSL_2(\mathbb{Z})$ and satisfying $\Gamma_\infty = B(\mathbb{Z})$. The following formula (cf. Theorem 3.3.2) encodes the relationship precisely.

$$R_\Gamma^{(-1)}(z) = \mathrm{e}(-z) + \frac{1}{2}\mathrm{c}_\Gamma(-1,0) + \sum_{n>0}\mathrm{c}_\Gamma(-1,n)\,\mathrm{e}(nz) \tag{1.2.10}$$

1.3 Conjugate Rademacher sums

It is striking that the constant term in the Fourier expansion (1.2.10) of the Rademacher sum $R_\Gamma^{(-1)}(z)$ is $\mathrm{c}_\Gamma(-1,0)/2$ and not $\mathrm{c}_\Gamma(-1,0)$, given that the coefficient of $\mathrm{e}(nz)$ in (1.2.10) is exactly $\mathrm{c}_\Gamma(-1,n)$ for $n > 0$. One may also observe that the formula (1.2.6) defining the values $\mathrm{c}_\Gamma(-1,n)$ makes sense for arbitrary integers $n \in \mathbb{Z}$, and so it is natural to ask what rôle is played by the $\mathrm{c}_\Gamma(-1,n)$ for $n < 0$? An answer to this question is obtained when we consider the *conjugate Rademacher sum associated to* Γ, denoted $\bar{R}_\Gamma^{(-1)}(z)$, and defined by setting

$$\bar{R}_\Gamma^{(-1)}(z) = \mathrm{e}(-\bar{z}) + \lim_{K\to\infty}\sum_{\gamma\in(B(\mathbb{Z})\backslash\Gamma)^\times_{\leq K}} \mathrm{e}(-\gamma\cdot\bar{z}) - \mathrm{e}(-\gamma\cdot\infty) \tag{1.3.1}$$

for $z \in \mathbb{H}$. The conjugate Rademacher sum $\bar{R}_\Gamma^{(-1)}(z)$ defines an anti-holomorphic function on $\mathbb{H}$. At first glance it appears that we should recover the classical Rademacher sum $R_\Gamma^{(-1)}(z)$ after substituting $\bar{z}$ for z in $\bar{R}_\Gamma^{(-1)}(z)$, since the expressions defining $R_\Gamma^{(-1)}(z)$ and $\bar{R}_\Gamma^{(-1)}(\bar{z})$ would appear to coincide, but in fact the delicate limit defining the Rademacher sums behaves very differently depending on which half plane the variable z lies in, and the difference $R_\Gamma^{(-1)}(z) - \bar{R}_\Gamma^{(-1)}(\bar{z})$ can be rather far from vanishing. As an illustration of this, we show in §3.3 that the Fourier expansion of the conjugate Rademacher sum $\bar{R}_\Gamma^{(-1)}(z)$ in terms of the anti-holomorphic exponential $\mathrm{e}(\bar{z})$ recovers the values $\mathrm{c}_\Gamma(-1,n)$ for $n < 0$ as Fourier coefficients. More precisely, we establish the following counterpart to (1.2.10) (cf. Theorem 3.3.3).

$$\bar{R}_\Gamma^{(-1)}(z) = \mathrm{e}(-\bar{z}) - \frac{1}{2}\mathrm{c}_\Gamma(-1,0) - \sum_{n<0}\mathrm{c}_\Gamma(-1,n)\,\mathrm{e}(n\bar{z}) \tag{1.3.2}$$

We demonstrate in §3.4 that the holomorphic function $\bar{R}_{\Gamma}^{(-1)}(\bar{z})$ is also an abelian integral (cf. Theorem 3.4.4). More than this, the difference

$$R_{\Gamma}^{(-1)}(z) - \bar{R}_{\Gamma}^{(-1)}(z) \tag{1.3.3}$$

is a Γ-invariant harmonic function on $\mathbb{H}$ (cf. Theorem 3.4.6).

The Γ-invariance of the function (1.3.3) entails a remarkable formula for $\bar{R}_{\Gamma}^{(-1)}(z)$ in the case that Γ has genus zero. For since the holomorphic Rademacher sum $R_{\Gamma}^{(-1)}(z)$ is Γ-invariant in the genus zero case, the anti-holomorphic function $\bar{R}_{\Gamma}^{(-1)}(z)$ must also be Γ-invariant. Given our knowledge (1.3.2) of the Fourier expansion of $\bar{R}_{\Gamma}^{(-1)}(z)$ it follows that $\bar{R}_{\Gamma}^{(-1)}(z)$ must be identically constant, and furthermore, this constant must be $-\mathrm{c}_{\Gamma}(-1,0)/2$. In the case that $\Gamma = PSL_2(\mathbb{Z})$ we have $\mathrm{c}_{\Gamma}(-1,0) = 24$, and we thus obtain the remarkable identity

$$\mathrm{e}(-\bar{z}) + \lim_{K\to\infty} \sum_{\substack{0<c<K \\ -K^2<d<K^2 \\ (c,d)=1}} \mathrm{e}\left(-\frac{a\bar{z}+b}{c\bar{z}+d}\right) - \mathrm{e}\left(-\frac{a}{c}\right) = -12, \tag{1.3.4}$$

where in each summand a and b are integers chosen so that $ad - bc = 1$. This implies that the function $R_{\Gamma}^{(-1)}(z) - \bar{R}_{\Gamma}^{(-1)}(z)$ is a (non-normalized) hauptmodul for Γ when Γ has genus zero, and in particular,

$$\mathrm{c}_{\Gamma}(-1,n) = \delta_{-1,n} \tag{1.3.5}$$

for $n < 0$. We recover the normalized hauptmodul for Γ (cf. §2.5) by considering the sum $R_{\Gamma}^{(-1)}(z) + \bar{R}_{\Gamma}^{(-1)}(z)$.

1.4 Solid tori

It was observed in [CN79] that each group attached to the Monster via monstrous moonshine may be described as a group of *$n\|h$-type*, meaning a discrete group $\Gamma < PSL_2(\mathbb{R})$ of the form $\Gamma = \Gamma_0(n\|h) + S$ for some $n, h \in \mathbb{Z}^+$ with $h|(n,24)$, and S a subgroup of the group of exact divisors of n/h. For such n, h, and S, the group $\Gamma_0(n\|h)+S$ contains and normalizes $\Gamma_0(nh)$, and has the property that $(\Gamma_0(n\|h)+S)_\infty = B(\mathbb{Z})$. (We recall the precise definition of $\Gamma_0(n\|h)+S$ in §2.5.)

We shed new light on the groups of $n\|h$-type by demonstrating in §6.2 that they solve a natural family of moduli problems for elliptic curves equipped with certain kinds of extra structure. As we will explain presently, a common feature of all the members of this family is a choice of filling of the elliptic curve in question; that is, a solid torus whose boundary is the given elliptic curve. Hence, we interpret the $n\|h$-type groups as defining moduli of (decorated) solid tori.

The fact that each group Γ of $n\|h$-type satisfies $\Gamma_\infty = B(\mathbb{Z})$ suggests the importance of the quotient $B(\mathbb{Z})\backslash\mathbb{H}$. Regarding $\mathbb{H}$ as a moduli space for triples (E,γ,γ') where E is an elliptic curve and (γ,γ') is an oriented basis for the first homology group $H_1(E,\mathbb{Z})$ of E, we see that $B(\mathbb{Z})\backslash\mathbb{H}$ parameterizes pairs (E,γ) where E is an elliptic curve and γ is a primitive element of $H_1(E,\mathbb{Z})$. To each such pair (E,γ) is naturally associated an infinite volume hyperbolic 3-manifold with boundary an elliptic curve, for the cycle γ specifies a way to fill in the surface defined by E, thus yielding a solid torus. According to a theorem of Sullivan (cf. [Sul81], [McM92]) every complete, smooth, infinite volume hyperbolic 3-manifold with boundary an elliptic curve arises in this fashion, so we consider pairs (E,C) where E is an elliptic curve over $\mathbb{C}$ and C is an oriented subgroup of E isomorphic to S^1. We call such a pair a *solid torus*; according to our discussion the quotient space $B(\mathbb{Z})\backslash\mathbb{H}$ naturally parameterizes solid tori.

For (E,C) a solid torus and $n\in\mathbb{Z}^+$ we write $C[n]$ for the group of n-division points of C, and we define an *n-compatible isogeny of solid tori* $(E',C')\to(E,C)$ to be an isogeny $E'\to E$ of elliptic curves that maps $C'[n]$ to a subgroup of $C[n]$, and we define a *fully compatible isogeny of solid tori* to be an isogeny $E'\to E$ of elliptic curves that restricts to an orientation preserving map $C'\to C$ on the underlying primitive cycles. (Note that the n in an n-compatible isogeny of solid tori does not restrict the degree of the underlying isogeny of elliptic curves.) For (E,C) a solid torus, the canonical map $E\to E/C[n]$ may be viewed as defining a fully compatible isogeny of solid tori $(E,C)\to(E,C)/C[n]$, where $(E,C)/C[n]$ is a shorthand for the solid torus whose underlying elliptic curve is $E/C[n]$, and whose primitive cycle is the image of C under the natural map $E\to E/C[n]$. It may happen that an n-compatible isogeny induces invertible n/e-compatible isogenies (i.e. *n/e-compatible isomorphisms*) $(E',C')/C'[e]\to(E,C)$ and $(E',C')\to(E,C)/C[e]$ for some divisor e of n. In this case we say that (E',C') and (E,C) are *$n+e$-related*. By considering the existence or otherwise of such morphisms of solid tori we arrive at moduli interpretations for all the $n\|h$-type groups. For example, given a subgroup S of the group of exact divisors of n, say solid tori (E,C) and (E',C') are *$n+S$-related* if (E,C) and (E',C') are $n+e$-related for some $e\in S$. Then the $n+S$-relation is an equivalence relation on solid tori, and the quotient $(\Gamma_0(n)+S)\backslash\mathbb{H}$ is in natural correspondence with $n+S$-equivalence classes of solid tori.

In this way we demonstrate that the groups of $n\|h$-type may be explained as being the groups that solve a natural family of moduli problems for decorated solid tori with conformal structure on the boundary.

1.5 Moonshine via Rademacher sums

The fact that the function (1.2.5) recovers the normalized hauptmodul of the group Γ when Γ has genus zero is a strong indication that the Rademacher sum may be useful for the purpose of understanding the functions of monstrous moonshine. A group theoretic characterization of the functions arising as McKay–Thompson series was given in [CMS04]. There the authors showed that a holomorphic function on $\mathbb{H}$ coincides with $T_g(z)$ for some $g \in \mathbb{M}$ if and only if it is the normalized hauptmodul for a group Γ satisfying each of the following properties.

1. The Riemann surface $\Gamma\backslash\mathbb{H}$ has genus zero.
2. The group Γ is of $n\|h$-type.
3. The group Γ is an extension of $\Gamma_0(nh)$ by a group of exponent 2.
4. Each cusp of Γ can be mapped to the infinite cusp by an element $\sigma \in PSL_2(\mathbb{R})$ with the property that $(\sigma\Gamma\sigma^{-1})_\infty = B(\mathbb{Z})$ and the intersection $\Gamma \cap \sigma\Gamma\sigma^{-1}$ contains $\Gamma_0(nh)$.

As we saw in §1.2 the variance of the Rademacher sum $R_\Gamma^{(-1)}(z)$ with respect to the action of Γ detects whether or not Γ has genus zero, and we have discussed in §1.4 the significance of the groups appearing in condition 2. We show in §§6.3,6.4 that the conditions 3 and 4 also admit natural reformulations in terms of the Rademacher sums.

Our reformulations come into view when we generalize the construction (1.2.2) so as to associate a Rademacher sum $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{(-1)}(z)$ to each triple $(\Gamma, \mathsf{p}, \mathsf{q})$ where Γ is a group commensurable with $PSL_2(\mathbb{Z})$, and p and q are cusps of Γ. (This construction is presented in §3.) We write $R_{\Gamma,\mathsf{p}}^{(-1)}(z)$ as a shorthand for $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{(-1)}(z)$ when q is the *infinite cusp* $\Gamma \cdot \infty$. It develops in §3.4 that $R_{\Gamma,\mathsf{p}}^{(-1)}(z)$ is an abelian integral for Γ with a simple pole at p, and no other singularities, and the expansion of $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{(-1)}(z)$ at the infinite cusp coincides with that of $R_{\Gamma,\mathsf{p}}^{(-1)}(z)$ at q, up to (addition by) a certain constant function (cf. Theorem 3.4.4). Considering all the functions $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{(-1)}(z)$ for varying cusps p and q we are able to access more subtle properties of the curve $\mathsf{X}_\Gamma = \Gamma\backslash\mathbb{H} \cup \Gamma\backslash\hat{\mathbb{Q}}$ (cf. §2.5) beyond its genus. We find in §6.3 that condition 3 translates into a certain condition of symmetry (cf. Proposition 6.3.2) in the Rademacher sums associated to the Hecke congruence group $\Gamma_0(nh)$, where n and h are as in condition 2, and we find in §6.4 that condition 4 translates into the condition (cf. Proposition 6.4.3) that the function $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{(-1)}$ be invariant for this subgroup $\Gamma_0(nh)$.

Our moduli interpretation of the $n\|h$-type groups, together with the invariance and symmetry conditions on Rademacher sums just described, facilitate a reformulation of the conditions of [CMS04]. We thus arrive at a new characterization (cf. Theorem 6.5.1) of the functions of monstrous moonshine in terms of Rademacher sums, and moduli of solid tori.

1.6 Modified Rademacher sums

In order to understand better the behavior of the Rademacher sums $R_\Gamma^{(-1)}(z)$, and, in particular, the subtraction of the constant $\mathrm{e}(-\gamma \cdot \infty)$ in (1.2.2), we consider (also in §3) a generalization of the Rademacher sum adapted to the problem of constructing modular forms of arbitrary even integer weight. Our approach to Rademacher sums of negative weight is inspired by the work of Niebur in [Nie74].

For $\kappa, m \in \mathbb{Z}$ we define the *Rademacher sum of weight* 2κ *and order* m *associated to* Γ by setting

$$R_\Gamma^{\kappa(m)}(z) = \mathrm{e}(mz) + \lim_{K\to\infty} \sum_{\gamma\in(B(\mathbb{Z})\backslash\Gamma)^\times_{\leq K}} \mathrm{e}(m\gamma\cdot z)\,\mathrm{r}^\kappa(m,\gamma,z)(cz+d)^{-2\kappa}, \tag{1.6.1}$$

where $c, d \in \mathbb{R}$ are chosen (for each summand) so that γ is the image of a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{R})$ in $PSL_2(\mathbb{R})$, and $\mathrm{r}^\kappa(m,\gamma,z)$ is the *Rademacher regularization factor of weight* 2κ, satisfying

$$\mathrm{r}^\kappa(m,\gamma,z) = 1 - \mathrm{e}(m\gamma\cdot\infty - m\gamma\cdot z)\,\mathrm{e}(m\gamma\cdot z - m\gamma\cdot\infty)_{<1-2\kappa}, \tag{1.6.2}$$

where $\mathrm{e}(z)_{<K}$ denotes the *partial exponential function* $\mathrm{e}(z)_{<K} = \sum_{0\leq k<K}(2\pi\mathbf{i}z)^k/k!$ (cf. §3.1). Observe that we recover the Rademacher sum $R_\Gamma^{(-1)}(z)$ of (1.2.2) when $\kappa = 0$. Also, we have $\mathrm{r}^\kappa(m,\gamma,z) = 1$ when κ is positive, so that the Rademacher sum $R_\Gamma^{\kappa(m)}(z)$ is a holomorphic Poincaré series (cf. §3.1) when $\kappa > 0$ (which is absolutely and locally uniformly convergent for $\kappa > 1$).

In §3.1 we furnish generalizations of the functions $\mathrm{Kl}_\gamma(m,n)$ and $\mathrm{Bl}_\gamma(m,n)$ for arbitrary $\kappa \in \mathbb{Z}$, obtaining a generalization of the formula (1.2.6), and leading to the following analogue (cf. Theorem 3.3.2) of (1.2.10).

$$R_\Gamma^{\kappa(m)}(z) = \mathrm{e}(mz) + \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0) + \sum_{n>0}\mathrm{c}_\Gamma^\kappa(m,n)\,\mathrm{e}(nz) \tag{1.6.3}$$

The *conjugate Rademacher sum of weight* 2κ *and order* m *associated to* Γ is defined in direct analogy with the definition (1.3.1) of the conjugate Rademacher sum $\bar{R}_\Gamma^{(-1)}(z)$, which we presently recognize as the conjugate Rademacher sum of weight 0 and order 1 associated to Γ.

$$\bar{R}_\Gamma^{\kappa(m)}(z) = \mathrm{e}(m\bar{z}) + \lim_{K\to\infty} \sum_{\gamma\in(B(\mathbb{Z})\backslash\Gamma)^\times_{\leq K}} \mathrm{e}(m\gamma\cdot\bar{z})\,\mathrm{r}^\kappa(m,\gamma,\bar{z})(c\bar{z}+d)^{-2\kappa}. \tag{1.6.4}$$

We have the following analogue of (1.3.2), and counterpart to (1.6.3), which leads naturally to analogues of (1.3.3) for arbitrary $\kappa \in \mathbb{Z}$ (cf. Theorem 3.3.3).

$$\bar{R}_\Gamma^{\kappa(m)}(z) = \mathrm{e}(m\bar{z}) - \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0) - \sum_{n<0}\mathrm{c}_\Gamma^\kappa(m,n)\,\mathrm{e}(n\bar{z}) \tag{1.6.5}$$

The variance of the functions $R_\Gamma^{\kappa(m)}(z)$ with respect to the natural (weight 2κ) action of the group Γ was described by Niebur in [Nie74] for the case that Γ has a single cusp. We verify in §4.5 that for Γ an arbitrary group commensurable with the modular group $PSL_2(\mathbb{Z})$, the function $R_\Gamma^{\kappa(m)}(z) + \mathrm{c}_\Gamma^\kappa(m,0)/2$ defines an *automorphic integral of weight* 2κ *for* Γ (cf. §2.7) when $m < 0$, meaning that $R_\Gamma^{\kappa(m)}(z)$ is holomorphic on $\mathbb{H}$, possibly having poles at cusps of Γ, and we have

$$\left(R_\Gamma^{\kappa(m)}(\gamma \cdot z) + \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0)\right)(cz+d)^{-2\kappa} = R_\Gamma^{\kappa(m)}(z) + \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0) + \omega(\gamma)(z) \tag{1.6.6}$$

for some function $\omega : \Gamma \to \mathbb{C}[z]$, where the polynomial $\omega(\gamma)(z)$ has degree at most -2κ in z (cf. Theorem 4.5.1, Proposition 4.4.10). An identification of the function ω may be given in terms of a certain canonically defined map $I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma)$, where $I_\kappa(\Gamma)$ denotes the space of automorphic integrals of weight 2κ for Γ (cf. §2.7), and $S_\kappa(\Gamma)$ denotes the space of cusp forms of weight 2κ for Γ. We show (cf. Theorem 4.5.4) that the sequence

$$0 \to M_\kappa(\Gamma) \to I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma) \to 0 \tag{1.6.7}$$

is exact, where $M_\kappa(\Gamma)$ denotes the space of modular forms of weight 2κ for Γ possibly having poles at cusps (cf. §2.7), and the map $M_\kappa(\Gamma) \to I_\kappa(\Gamma)$ is the natural inclusion.

Apart from the delicate convergence of the limit defining the Rademacher sum $R_\Gamma^{\kappa(m)}(z)$, the most curious feature of the functions $R_\Gamma^{\kappa(m)}(z)$ is that they transform naturally with respect to the group Γ only after the addition of the constant $\mathrm{c}_\Gamma^\kappa(m,0)/2$, which is typically non-zero. We overcome this feature of the classical Rademacher sum by introducing in §4 the *continued Rademacher sum of weight* 2κ *and order* m *associated to* Γ, defined for $\kappa \leq 0$ and $m < 0$ by setting

$$T_\Gamma^{\kappa(m)}(z,s) = \mathrm{e}(mz) + \sum_{\gamma \in (B(\mathbb{Z})\backslash\Gamma)^\times} \mathrm{e}(m\gamma \cdot z)\,\mathrm{t}^\kappa(m,\gamma,z,s)(cz+d)^{-2\kappa} \tag{1.6.8}$$

where $\mathrm{t}^\kappa(m,\gamma,z,s)$ is a generalization of the Rademacher regularization factor $\mathrm{r}^\kappa(m,\gamma,z)$ of (1.6.2) satisfying $\mathrm{t}^\kappa(m,\gamma,z,1) = \mathrm{r}^\kappa(m,\gamma,z)$.

The right hand side of (1.6.8) converges absolutely and locally uniformly in z and s for $z \in \mathbb{H}$ and $\Re(s) > 1$. We define $T_\Gamma^{\kappa(m)}(z)$ by taking the limit as s tends to 1 in $T_\Gamma^{\kappa(m)}(z,s)$. We find that

$$T_\Gamma^{\kappa(m)}(z) = \lim_{s\to 1^+} T_\Gamma^{\kappa(m)}(z,s) = R_\Gamma^{\kappa(m)}(z) - \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0) \tag{1.6.9}$$

(cf. Proposition 4.4.9), so that the Fourier expansion of $T_\Gamma^{\kappa(m)}(z)$ has vanishing constant term. Then, in the case that Γ has genus zero, the function $T_\Gamma^{(-1)}(z) = T_\Gamma^{0(1)}(z)$ is precisely the normalized hauptmodul of Γ. We call $T_\Gamma^{\kappa(m)}(z)$ the *normalized Rademacher sum of weight* 2κ *and order* m *associated to* Γ. We may regard the disappearance of the constant term in (1.6.9) as a consequence of the noncommutativity of the limits $s \to 1$ and $K \to \infty$ when applied to (the expression obtained by replacing $(B(\mathbb{Z})\backslash\Gamma)^\times$ with $(B(\mathbb{Z})\backslash\Gamma)^\times_{\leq K}$ in) the right hand side of (1.6.8).

To recover functions with natural modular properties at non-positive weights we modify the continued Rademacher sum by subtracting a renormalized value at $z = 0$.

$$Q_\Gamma^{\kappa(m)}(z,s) = T_\Gamma^{\kappa(m)}(z,s) - T_{0,\Gamma}^{\kappa(m)}(s) \tag{1.6.10}$$

The function $T_{0,\Gamma}^{\kappa(m)}(s)$ is defined by an expression analogous to (1.6.8) which also converges absolutely and locally uniformly in s for $\Re(s) > 1$, and would vanish at $s = 1$ if it were not for the noncommutativity of limits as $s \to 1$ and $K \to \infty$. We define the *modified Rademacher sum of weight* 2κ *and order* m *associated to* Γ, denoted $Q_\Gamma^{\kappa(m)}(z)$, by taking the limit as s tends to 1 in $Q_\Gamma^{\kappa(m)}(z,s)$, and as a counterpart to (1.6.9) we have

$$Q_\Gamma^{\kappa(m)}(z) = \lim_{s\to 1^+} Q_\Gamma^{\kappa(m)}(z,s) = R_\Gamma^{\kappa(m)}(z) + \frac{1}{2}\mathrm{c}_\Gamma^\kappa(m,0) \tag{1.6.11}$$

(cf. Proposition 4.4.10), so that, in light of (1.6.6), the modified Rademacher sum $Q_\Gamma^{\kappa(m)}(z)$ is an automorphic integral of weight 2κ for Γ.

The definition of the modified Rademacher sum $Q_\Gamma^{\kappa(m)}(z)$ is inspired by Hurwitz's relation for the Hurwitz zeta function, and our continuation procedure may be regarded as identifying the factor $1/2$, appearing in the constant term of the Fourier expansion (1.6.3) of the classical Rademacher sum of non-positive weight, with -1 times the value of the Riemann zeta function $\zeta(s)$ at $s = 0$ (cf. §4.3).

In order to obtain and analyze spanning sets for spaces of automorphic integrals we attach modified Rademacher sums $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ to each triple $(\Gamma, \mathsf{p}, \mathsf{q})$ where Γ is a group commensurable with the modular group $PSL_2(\mathbb{Z})$ and p and q are cusps of Γ (cf. §4.1). The function $Q_\Gamma^{\kappa(m)}(z)$ is then recovered upon taking both p and q to be the *infinite cusp* $\Gamma\cdot\infty$. If we write $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ as a shorthand for $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ when $\mathsf{q} = \Gamma\cdot\infty$, then, as is shown in §4.5, the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$, for varying cusps $\mathsf{p} \in \Gamma\backslash\hat{\mathbb{Q}}$ and positive integers $m \in \mathbb{Z}^+$, constitute a basis for the space $I_\kappa(\Gamma)$ when $\kappa < 0$ (cf. Theorem 4.5.3). In the case that $\kappa = 0$ the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ span a subspace of $I_0(\Gamma)$ of codimension 1, and a full basis is obtained by including a constant function. The function $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ defines an automorphic integral with a single pole at the cusp p, and the function $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ encodes the Fourier expansion of $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ at the cusp q, and no correction of constant terms is necessary for the validity of these statements.

1.7 Hecke operators

The original result of Rademacher [Rad39] identifies the first Rademacher sum $R_\Gamma^{(-1)}(z)$, for $\Gamma = PSL_2(\mathbb{Z})$, with the function $J(z) + 12$, where $J(z)$ is the normalized hauptmodul for the modular group. In light of this it is natural to try to identify more general Rademacher sums in a similar fashion. In this section we will take $\kappa = 0$, but we will allow m to be an arbitrary non-zero integer. The identity

$$R_\Gamma^{(m)}(z) = \overline{\bar{R}_\Gamma^{(-m)}(z)} \tag{1.7.1}$$

follows immediately from the definitions of the classical and conjugate Rademacher sums, demonstrating that the conjugate Rademacher sums of negative order may be recovered from the classical Rademacher sums with positive order. On the other hand, one can show, as we did for the special case that $m = -1$, that $\bar{R}_\Gamma^{(m)}(z)$ is the constant function with constant value $-\mathrm{c}_\Gamma(m, 0)/2$ in case $m < 0$ and Γ has genus zero. When $\Gamma = PSL_2(\mathbb{Z})$ we have $\mathrm{c}_\Gamma(m, 0) = 24\sigma(-m, 1)$ for $m < 0$, where $\sigma(n, 1)$ is the sum of the divisors of n, so we have the following generalization of (1.3.4) for $m < 0$.

$$\mathrm{e}(m\bar{z}) + \lim_{K\to\infty} \sum_{\substack{0<c<K \\ -K^2<d<K^2 \\ (c,d)=1}} \mathrm{e}\left(m\frac{a\bar{z}+b}{c\bar{z}+d}\right) - \mathrm{e}\left(m\frac{a}{c}\right) = -12\sigma(-m, 1) \tag{1.7.2}$$

To obtain an expression for the classical Rademacher sums $R_\Gamma^{(m)}(z)$ with negative m we can analyze again its Fourier coefficients $\mathrm{c}_\Gamma(m, n)$. It follows from our results in §6.1 that

$$\mathrm{c}_\Gamma(m, n) = \delta_{m,n} \tag{1.7.3}$$

for $m, n < 0$ in case Γ has genus zero. The resulting function $T_\Gamma^{(m)}(z) = R_\Gamma^{(m)}(z) - \mathrm{c}_\Gamma(m, 0)/2$, satisfying the identity

$$T_\Gamma^{(m)}(z) + \frac{1}{2}\mathrm{c}_\Gamma(m, 0) = \mathrm{e}(mz) + \lim_{K\to\infty} \sum_{\gamma\in(B(\mathbb{Z})\backslash\Gamma)^\times_{\leq K}} \mathrm{e}(m\gamma\cdot z) - \mathrm{e}(m\gamma\cdot\infty), \tag{1.7.4}$$

is therefore holomorphic on the upper half plane, invariant for the action of Γ, and of the form $q^{-|m|} + o(1)$ for $q = \mathrm{e}(z)$. Consequently each normalized Rademacher sum $T_\Gamma^{(-m)}(z)$, for $m > 0$, is expressible as a degree m polynomial in $T_\Gamma^{(-1)}(z) = T_\Gamma(z)$. This polynomial is none other than the so-called *m-th Faber polynomial for* $T_\Gamma(z)$.

Proceeding from another direction, we can obtain the functions $T_\Gamma^{(-m)}(z)$ from the normalized hauptmodul $T_\Gamma(z)$ by applying Hecke operators. In §5.5 we study, in particular, the case that Γ is the modular group $PSL_2(\mathbb{Z})$. Recall that for $n \in \mathbb{Z}^+$ the action of the Hecke operator $\hat{T}(n)$

on a modular function $f(z)$ may be given by setting

$$(\hat{T}(n)f)(z) = \frac{1}{n} \sum_{\substack{ad=n \\ 0\leq b<d}} f\left(\frac{az+b}{d}\right). \tag{1.7.5}$$

From this description one may deduce that the difference $T_\Gamma^{(-m)}(z) - m(\hat{T}(m)T_\Gamma)(z)$ is holomorphic in $\mathbb{H}$ and vanishes as $z \to \mathbf{i}\infty$. We thus obtain the remarkable expression

$$T_\Gamma^{(-m)}(z) = m(\hat{T}(m)T_\Gamma)(z) \tag{1.7.6}$$

for the higher order Rademacher sums which generalizes the original result of Rademacher. Combining the presentation (1.2.1) of $T_\Gamma(z)$ as a Rademacher sum with the definition (1.7.5) of the Hecke operator $\hat{T}(m)$ we obtain the expression

$$m(\hat{T}(m)T_\Gamma)(z) + \frac{1}{2}\mathrm{c}_\Gamma(-m,0) = \mathrm{e}(-mz) + \lim_{K\to\infty} \sum_{\gamma\in(B(\mathbb{Z})\backslash M(m))^\times_{\leq K}} \mathrm{e}(-\gamma\cdot z) - \mathrm{e}(-\gamma\cdot\infty), \tag{1.7.7}$$

where $M(m)$ denotes the (image in $PGL_2^+(\mathbb{Q})$ of the) set of 2×2 matrices with integral entries and determinant m. Comparison of (1.7.4) with (1.7.7), in view of the identity (1.7.6), suggests an equality of some, if not all, exponential terms in both sums, and in fact there is an injective map

$$\begin{aligned} B(\mathbb{Z})\backslash M(1) &\hookrightarrow B(\mathbb{Z})\backslash M(m) \\ \gamma &\mapsto \tilde{\gamma}, \end{aligned} \tag{1.7.8}$$

since $M(1) = \Gamma$ when $\Gamma = PSL_2(\mathbb{Z})$, with the property that $\mathrm{e}(-m\gamma\cdot z) = \mathrm{e}(-\tilde{\gamma}\cdot z)$. In the case that m is prime, the remaining terms assemble into a fractional power Rademacher sum

$$\mathrm{e}\left(-\frac{1}{m}z\right) + \lim_{K\to\infty} \sum_{\gamma\in(B(m\mathbb{Z})\backslash\Gamma)^\times_{\leq K}} \mathrm{e}\left(-\frac{1}{m}\gamma\cdot z\right) - \mathrm{e}\left(-\frac{1}{m}\gamma\cdot\infty\right), \tag{1.7.9}$$

where $B(m\mathbb{Z})$ is the *m-fold translation group*, generated by $z \mapsto z+m$. Thus we obtain another injective correspondence

$$\begin{aligned} B(m\mathbb{Z})\backslash M(1) &\hookrightarrow B(\mathbb{Z})\backslash M(m) \\ \gamma &\mapsto \tilde{\gamma} \end{aligned} \tag{1.7.10}$$

with the property that $\mathrm{e}(-\gamma\cdot z/m) = \mathrm{e}(-\tilde{\gamma}\cdot z)$. The identity (1.7.7) implies the vanishing of the fractional power Rademacher sum (1.7.9). This can also be proven directly, and we give a general vanishing result for fractional Rademacher sums in §5.4. More generally we obtain fractional power sums of the form

$$\mathrm{e}\left(-\frac{m}{l^2}z\right)+\lim_{K\to\infty}\sum_{\gamma\in(B(l\mathbb{Z})\backslash\Gamma)^{\times}_{\leq K}}\mathrm{e}\left(-\frac{m}{l^2}\gamma\cdot z\right)-\mathrm{e}\left(-\frac{m}{l^2}\gamma\cdot\infty\right), \tag{1.7.11}$$

for each exact divisor l of m, and each of these fractional power sums vanishes except for the sum corresponding to $l=1$. Thus our analysis covers generalizations of the Rademacher sums to fractional orders in addition to the higher integral orders. All the results discussed in this section admit generalizations to arbitrary groups commensurable with the modular group but we restrict ourselves to the modular group in this article.

1.8 Monstrous Lie algebras

The presentation (1.7.6) of the functions $T_\Gamma^{(-m)}(z)$ by means of Hecke operators (we restrict to the case $\Gamma=PSL_2(\mathbb{Z})$ in §1.7) immediately implies that their Fourier coefficients are positive integers. This fact points to the existence of further algebraic structures beyond the moonshine module vertex operator algebra $V^\natural$. In his proof of the moonshine conjectures [Bor92] Borcherds introduced the Monster Lie algebra $\mathfrak{m}$, which admits a presentation as a bi-graded generalized Kac–Moody algebra. The key to Borcherds' method is the denominator identity for $\mathfrak{m}$. This in turn can be interpreted as a BGG-type resolution of the trivial $\mathfrak{m}$-module. Let $\mathcal{V}$ denote the Verma module with highest weight 0 for $\mathfrak{m}$. Then $\mathcal{V}$ is the first term in this BGG-type resolution of the trivial module, and the bi-graded dimension $\operatorname{gdim}\mathcal{V}$ of $\mathcal{V}$ is obtained by computing the coefficients of p and q in the expression

$$\operatorname{gdim}\mathcal{V}=\exp\left(\sum_{m>0}(\hat{T}(m)T_\Gamma)(z)p^m\right) \tag{1.8.1}$$

where $q=\mathrm{e}(z)$. Applying the operator $F\mapsto p\partial_p\log F$ to (1.8.1) we obtain a generating function for the higher (absolute) order normalized Rademacher sums $T_\Gamma^{(-m)}(z)$, by (1.7.6). It is well-known (cf. [Mac95]) that this operator relates the complete homogeneous symmetric functions to the power-sum symmetric functions. Thus the bi-graded dimension of the Verma module for the Monster Lie algebra may be viewed as the *complete Rademacher sum* (of weight 0) associated to Γ.

In §7.2 we consider a family of generalized Kac–Moody algebras $\{\mathfrak{m}_g\}$, indexed by elements g in the Monster group $\mathbb{M}$. Just as the Monster Lie algebra may be constructed from the moonshine module $V^\natural$, the monstrous Lie algebras $\mathfrak{m}_g$, studied by Carnahan in [Car07], may be constructed from the g-twisted $V^\natural$-modules $V^\natural_g$, for $g\in\mathbb{M}$. The graded dimensions of the spaces

$V_g^\natural$ are given by functions $J_g(z) = J_{\Gamma_g}(z)$, related to the McKay–Thompson series $T_g(z)$ via the involution $z \mapsto -1/z$.

$$J_g(z) = T_g(-1/z) \tag{1.8.2}$$

One should note that Borcherds set the precedent here, introducing a family $\{\mathfrak{m}'_g\}$ of monstrous Lie superalgebras in [Bor92]. These algebras $\mathfrak{m}'_g$ are more directly related to the functions $T_g(z)$, rather than the $J_g(z)$, and for our purposes the algebras $\mathfrak{m}_g$ appear to be more convenient.

The monstrous Lie algebras $\mathfrak{m}_g$ are constructed in [Car07] using the semi-infinite cohomology version of the no-ghost theorem. We identify the bi-graded subspaces of the algebra $\mathfrak{m}_g$ in terms of the twisted modules $V_h^\natural$, with $h \in \langle g \rangle$. One can show that the $\mathfrak{m}_g$ are generalized Kac–Moody algebras, and one can also deduce formulas for the bi-graded dimensions of their Verma modules $\mathcal{V}_g$, generalizing (1.8.1).

$$\operatorname{gdim} \mathcal{V}_g = \exp\left(\sum_{m>0} (\hat{T}(m)T_g)(z)p^m \right) \tag{1.8.3}$$

Carnahan [Car07] also found a remarkable denominator identity which yields an alternative expression for the graded dimension of $\mathcal{V}_g$; namely,

$$\operatorname{gdim} \mathcal{V}_g = \frac{1}{p(T_g(w) - J_g(z))} \tag{1.8.4}$$

where $p = \mathrm{e}(w)$. This identity (1.8.4) demonstrates that the bi-graded dimension can be viewed as a meromorphic function on $\mathbb{H} \times \mathbb{H}$ with poles at $w \in \Gamma_g \cdot (-1/z)$.

We also show in §7.2 that the bi-graded dimension of $\mathcal{V}_g$ is bounded below, coefficient-wise, by the generating function

$$\sum_{m>0} J_g^{(-m)}(z)p^m \tag{1.8.5}$$

where $J_g^{(-m)}(z) = T_g^{(-m)}(-1/z)$ when g is of Fricke type (cf. Proposition 7.2.4). This suggests that the Verma modules $\mathcal{V}_g$ may contain naturally defined subspaces whose bi-graded dimensions coincide with the expressions (1.8.5).

1.9 Chiral gravity

Our results on Rademacher sums, at first glance, add as much to the mystery of monstrous moonshine as they reveal. We have formulated a characterization of the groups of monstrous moonshine in terms of Rademacher sums and solid tori. Having done so we face the new question of where the Rademacher sums themselves appear in relation to the moonshine module vertex operator algebra $V^\natural$? It appears now that the relation to physics should again be useful, as it was in the construction of $V^\natural$ [FLM88], and the proof of the moonshine conjecture [Bor92].

In §7.1 we explain how the Rademacher sums associated to elements of the Monster might be regarded as giving strong evidence for the existence of another construction of the moonshine module $V^\natural$, and its twisted sectors $V_g^\natural$, in which all the features of monstrous moonshine, including the genus zero property, become transparent. In fact the original Rademacher sum (1.2.1) does appear as a saddle point approximation to the partition function of the simplest chiral three dimensional quantum gravity. (See [Wit07], [MW07], [Man07], [MM07], [LSS08] and [MSS09] for the development of this idea in the physics literature.) Our new identity (1.3.4) for the conjugate Rademacher sum must reflect the chiral nature of this three dimensional quantum gravity.

We expect that our analytic continuation of the Rademacher sums, and the explanation of the appearance of the constant term, will also develop from a rigorous analysis of the saddle point approximation using a zeta function regularization of the chiral gravity partition function, as was done for various models of two dimensional CFT. Our analysis of the Rademacher sums corresponding to the discrete groups of monstrous moonshine, and of moduli spaces of decorated solid tori points to a description of all the spaces $V_g^\natural$, for $g \in \mathbb{M}$, via a family of g-twisted versions of the simplest chiral three dimensional quantum gravity. We expect that our reformulation of monstrous moonshine in terms of Rademacher sums and moduli of solid tori will eventually be understood from the properties of this remarkable chiral three dimensional quantum gravity.

In §7.3 we also elucidate the relationship between the higher order Rademacher sums, and the key object of Borcherds' proof of the moonshine conjecture; viz., the Monster Lie algebra, and its twisted counterparts. This leads us to the stringy quantization (cf. [DMVV97]) of $V^\natural$, and we generalize this to all twisted sectors $V_g^\natural$, for $g \in \mathbb{M}$. We prove in Theorem 7.3.1 that the resulting second quantized spaces have the same bi-graded dimensions as the Verma modules $\mathcal{V}_g$ (cf. (1.8.3)) and thus may be viewed as stringy realizations of the representations of the generalized Kac–Moody algebras $\mathfrak{m}_g$. It is an interesting problem to interpret the spaces $\mathcal{V}_g$, and their subspaces with graded dimension given by the higher order normalized Rademacher sums (1.8.5), in terms of the second quantized chiral three dimensional quantum gravities. In particular, the simple description of the singularities of the partition function given by (1.8.4) should have a natural interpretation in terms of quantum gravity.

Thus our present results on the Rademacher sums, and the new conjectures about their origin, clearly indicate that the real nature of moonshine might not be as remote anymore, and more than this, the full examination of its structure might give new insight into the fundamental problem of modern physics.

2 Conventions

We write $\mathbb{Z}^+$ for the set of positive integers, and the notations $\mathbb{Q}^+$ and $\mathbb{R}^+$ are to be interpreted similarly. We write $\mathbb{N}$ for the set of non-negative integers. For R a ring without zero divisors, we write $R^\times$ for the multiplicative monoid of non-zero elements in R.

2.1 Functions

For $z, s \in \mathbb{C}$ with $z \neq 0$, we write z^s as a shorthand for $\exp(s\log(z))$, where log denotes the principal branch of the logarithm, so that

$$-\pi < \Im(\log(z)) \leq \pi, \tag{2.1.1}$$

and we write $z^{(s)}$ as a shorthand for $z^s/\Gamma(s+1)$, where $\Gamma(s)$ denotes the Gamma function. We adopt the convention of setting

$$\mathrm{e}(z) = \exp(2\pi\mathbf{i}z) = \sum_{k\geq 0}(2\pi\mathbf{i}z)^{(k)} \tag{2.1.2}$$

for $z \in \mathbb{C}$. We write $\Phi(a, b, z)$ for the analytic function on $\mathbb{C}^3$ defined by setting

$$\Phi(a, b, z) = \frac{{}_1F_1(a; b; 2\pi\mathbf{i}z)}{\Gamma(b)} = \sum_{k\geq 0}\frac{\Gamma(k+a)}{\Gamma(a)\Gamma(k+b)}(2\pi\mathbf{i}z)^{(k)} \tag{2.1.3}$$

where ${}_1F_1(a; b; x)$ denotes the confluent hypergeometric function (cf. [AS64, §13]). Kummer's transformations for the confluent hypergeometric function yield the following identity for Φ.

$$\Phi(a, b, z) = \mathrm{e}(z)\Phi(b-a, b, -z) \tag{2.1.4}$$

The exponential function $\mathrm{e}(z)$ is a solution to the differential equation

$$(z\partial_z + 1 - s)(\partial_z - 2\pi\mathbf{i})u(z) = 0, \tag{2.1.5}$$

which has a regular singular point at 0 and an irregular singularity at ∞. Another solution is furnished by the function $z \mapsto \Phi(1, 1+s, z)(2\pi\mathbf{i}z)^s$, which we denote also by $\mathrm{e}(z, s)$.

$$\mathrm{e}(z, s) = \Phi(1, 1+s, z)(2\pi\mathbf{i}z)^s = \sum_{k\geq 0}(2\pi\mathbf{i}z)^{(k+s)} \tag{2.1.6}$$

Observe that we have $\mathrm{e}(z, n) = \mathrm{e}(z) - \mathrm{e}(z)_{<n}$ for $n \in \mathbb{Z}$, where $\mathrm{e}(z)_{<K}$ denotes the *partial exponential* $\mathrm{e}(z)_{<K} = \sum_{0\leq k<K}(2\pi\mathbf{i}z)^{(k)}$. In particular, $\mathrm{e}(z, n) = \mathrm{e}(z)$ when $n \leq 0$.

2.2 Divisors

For $n \in \mathbb{Z}^+$ and $s \in \mathbb{C}$ we write $\sigma(n,s)$ for the *divisor function*

$$\sigma(n,s) = \sum_{d|n} d^s. \tag{2.2.1}$$

For $a, b \in \mathbb{Z}$ we write (a,b) for the greatest common divisor of a and b, and we generalize this notation by writing $((a,b)$ for the greatest positive integer c say, such that $c^2|a$ and $c|b$, and by writing (a^∞, b) for the largest divisor of b that divides some power of a.

$$((a,b) = \max\left\{c \in \mathbb{Z}^+ \mid c^2|a \text{ and } c|b\right\} \tag{2.2.2}$$

$$(a^\infty, b) = \lim_{n\to\infty} (a^n, b) \tag{2.2.3}$$

For $d, n \in \mathbb{Z}^+$ we write $d\|n$ in the case that $d|n$ and $(d, n/d) = 1$, and call such a d an *exact divisor of* n. We write $\mathrm{Ex}(n)$ for the set of exact divisors of n. Then $\mathrm{Ex}(n)$ becomes a group of exponent 2 when equipped with the product $(e,f) \mapsto ef/(e,f)^2$. Indeed, if we write $\Pi(n)$ for the set of primes that divide n then there is a naturally defined isomorphism $\mathrm{Ex}(n) \cong (\mathbb{Z}/2)^{\Pi(n)}$ which associates to an exact divisor $e \in \mathrm{Ex}(n)$ the indicator function $\chi_e : \Pi(n) \to \mathbb{Z}/2$, satisfying $\chi_e(p) = 1$ if $p|e$ and $\chi_e(p) = 0$ otherwise.

Let Π denote the set of all primes. Then the infinite product $(\mathbb{Z}/2)^\Pi$ comes equipped with a naturally defined projection $(\mathbb{Z}/2)^\Pi \to (\mathbb{Z}/2)^{\Pi(n)}$, and a naturally defined section for it $(\mathbb{Z}/2)^{\Pi(n)} \to (\mathbb{Z}/2)^\Pi$, for each $n \in \mathbb{Z}^+$. Consequently we have naturally defined morphisms of groups $(\mathbb{Z}/2)^\Pi \to \mathrm{Ex}(n)$ and $\mathrm{Ex}(n) \to (\mathbb{Z}/2)^\Pi$ for each n, and for any $m, n \in \mathbb{Z}^+$ we have the natural map $\mathrm{Ex}(m) \to \mathrm{Ex}(n)$ furnished by the composition $\mathrm{Ex}(m) \to (\mathbb{Z}/2)^\Pi \to \mathrm{Ex}(n)$. If m divides n then the map $\mathrm{Ex}(m) \to \mathrm{Ex}(n)$ is injective, the map $\mathrm{Ex}(n) \to \mathrm{Ex}(m)$ is surjective, and the composition $\mathrm{Ex}(m) \to \mathrm{Ex}(n) \to \mathrm{Ex}(m)$ is the identity on $\mathrm{Ex}(m)$.

2.3 Isometries

Let us write $G(\mathbb{R})$ for the simple real Lie group $PSL_2(\mathbb{R})$. The group $G(\mathbb{R})$ acts naturally, from the left, on the upper-half plane $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$, and is just the group of orientation preserving isometries of $\mathbb{H}$ when we equip it with the *hyperbolic measure* $\mathrm{d}\mu(z) = \mathrm{d}x\mathrm{d}y/y^2$, for $z = x + \mathbf{i}y$.

It is convenient to enlist matrices in $SL_2(\mathbb{R})$ for the purpose of specifying elements of $G(\mathbb{R})$. We write $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ for the image of a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{R})$ in $G(\mathbb{R})$. Analogously, we write $A \mapsto [A]$ for the canonical map $SL_2(\mathbb{R}) \to G(\mathbb{R})$. The action of $G(\mathbb{R})$ on $\mathbb{H}$ is now described explicitly by the formula

$$\gamma \cdot z = \frac{az+b}{cz+d} \tag{2.3.1}$$

for $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. The matrix $\theta = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in GL_2(\mathbb{R})$ induces an outer-automorphism of $G(\mathbb{R})$, which we call *conjugation on* $G(\mathbb{R})$, and which we denote $\gamma \mapsto \bar{\gamma}$. Explicitly, we have

$$\bar{\gamma} = \begin{bmatrix} -a & b \\ c & -d \end{bmatrix} \Longleftarrow \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{2.3.2}$$

For each matrix $A \in GL_2^+(\mathbb{Q})$ there is a unique $\mu \in \mathbb{R}^+$ for which $A' = \mu A$ belongs to $SL_2(\mathbb{R})$. The assignment $GL_2^+(\mathbb{Q}) \to G(\mathbb{R})$ given by $A \mapsto [A']$ then factors through $PGL_2^+(\mathbb{Q})$, and the resulting map $PGL_2^+(\mathbb{Q}) \to G(\mathbb{R})$ is an embedding of groups. We write $G(\mathbb{Q})$ for the copy of $PGL_2^+(\mathbb{Q})$ in $G(\mathbb{R})$ obtained in this way. We write $G(\mathbb{Z})$ for the *modular group* $PSL_2(\mathbb{Z})$.

Extending the notation introduced above, we write $[A]$ and $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ for the image in $G(\mathbb{Q}) < G(\mathbb{R})$ of a matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $GL_2^+(\mathbb{Q})$. Given $\mu \in \mathbb{Q}^+$ we write $[\mu]$ as a shorthand for $[A]$ when A is the diagonal matrix $\left(\begin{smallmatrix} \mu & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Then $[\mu] \cdot z = \mu z$ for $z \in \mathbb{H}$.

$$[\mu] = \begin{bmatrix} \mu & 0 \\ 0 & 1 \end{bmatrix} \tag{2.3.3}$$

Let $B(\mathbb{R})$ denote the Borel subgroup of $G(\mathbb{R})$ consisting of images of upper-triangular matrices in $SL_2(\mathbb{R})$, so that $B(\mathbb{R})$ is just the subgroup of $G(\mathbb{R})$ that fixes the distinguished point ∞ on the *pointed real projective line* $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. Let $B(\mathbb{Q})$ and $B(\mathbb{Z})$ denote the intersections $B(\mathbb{R}) \cap G(\mathbb{Q})$ and $B(\mathbb{R}) \cap G(\mathbb{Z})$, respectively. Then $B(\mathbb{Z})$ is the group generated by the *modular translation* $T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$. Set $\hat{\mathbb{Q}} = \mathbb{Q} \cup \{\infty\} \subset \hat{\mathbb{R}}$. Then $G(\mathbb{Q})$ acts naturally on $\hat{\mathbb{Q}}$, and $B(\mathbb{Q})$ is just the subgroup of $G(\mathbb{Q})$ that fixes the point ∞. For $\alpha \in \mathbb{Q}$ we write T^α for the element $\left[\begin{smallmatrix} 1 & \alpha \\ 0 & 1 \end{smallmatrix}\right] \in B(\mathbb{Q})$.

$$T^\alpha = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \tag{2.3.4}$$

Then the elements of $B(\mathbb{Q})$ of the form T^α for $\alpha \in \mathbb{Q}$ constitute the *unipotent subgroup of* $B(\mathbb{Q})$, which we denote $B_u(\mathbb{Q})$. Given $\alpha \in \mathbb{Q}^+$ we write $B(\alpha\mathbb{Z})$ for the subgroup of $B_u(\mathbb{Q})$ generated by T^α.

$$B(\alpha\mathbb{Z}) = \left\{ \begin{bmatrix} 1 & \alpha n \\ 0 & 1 \end{bmatrix} \mid n \in \mathbb{Z} \right\} \tag{2.3.5}$$

The elements of $B(\mathbb{Q})$ of the form $[\mu]$ for $\mu \in \mathbb{Q}^+$ constitute the *diagonal subgroup of* $B(\mathbb{Q})$, which we denote $B_d(\mathbb{Q})$. The group $B(\mathbb{Q})$ is naturally isomorphic to the semidirect product $B_u(\mathbb{Q}) \rtimes B_d(\mathbb{Q})$. In particular, $B_u(\mathbb{Q})$ is a normal subgroup of $B(\mathbb{Q})$ and for any $\chi \in B(\mathbb{Q})$ we have $\chi = T^{\alpha}[\mu]$ for some uniquely determined $\alpha \in \mathbb{Q}$ and $\mu \in \mathbb{Q}^+$.

An element $\gamma \in G(\mathbb{Q})$ will have many preimages in $GL_2^+(\mathbb{Q})$, but among these there will be exactly two that have integral entries with no common divisor. Either of these two matrices will be called a *preferred representative* for γ. We define functions $P\det : G(\mathbb{Q}) \to \mathbb{Z}^+$ and $c : G(\mathbb{Q}) \to \mathbb{N}$ and $d : G(\mathbb{Q}) \to \mathbb{N}$ by setting

$$P\det(\gamma) = ad - bc, \quad c(\gamma) = |c|, \quad d(\gamma) = |d|, \tag{2.3.6}$$

in case $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2^+(\mathbb{Q})$ is a preferred representative for γ. We call $P\det$ the *projective determinant.* The group $G(\mathbb{Z})$ is exactly the preimage of 1 with respect to the projective determinant. The projective determinant is not multiplicative, but we have $P\det(\gamma\sigma) = P\det(\sigma) = P\det(\sigma\gamma)$ for all $\sigma \in G(\mathbb{Q})$ in case $P\det(\gamma) = 1$, so for any $n \in \mathbb{Z}^+$ the preimage of n under $P\det$ admits commuting left and right actions of $G(\mathbb{Z})$.

Observe that $c(T\gamma) = c(\gamma T) = c(\gamma)$ and $d(T\gamma) = d(\gamma)$. Given $X \subset G(\mathbb{Q})$ we set $X_\infty = X \cap B(\mathbb{Q})$ and $X^\times = X - X_\infty$. Then $X_\infty = \{\chi \in X \mid c(\chi) = 0\}$. We define $X^{\times\times}$ to be the subset of $X^\times$ consisting of $\chi \in X$ such that both $c(\chi)$ and $d(\chi)$ are non-zero.

$$X_\infty = \{\chi \in X \mid c(\chi) = 0\}\,, \tag{2.3.7}$$

$$X^\times = \{\chi \in X \mid c(\chi) \neq 0\}\,, \quad X^{\times\times} = \{\chi \in X \mid c(\chi)d(\chi) \neq 0\} \tag{2.3.8}$$

For $\gamma \in G(\mathbb{Q})$ with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ a preferred representative, we have

$$\gamma \cdot z - \gamma \cdot z' = \frac{P\det(\gamma)}{(cz+d)(cz'+d)}(z - z') \tag{2.3.9}$$

for $z, z' \in \mathbb{H}$, so that the derivative of the function $z \mapsto \gamma \cdot z$, which we will denote $\mathrm{jac}(\gamma, z)$, is given by

$$\mathrm{jac}(\gamma, z) = \frac{P\det(\gamma)}{(cz+d)^2}. \tag{2.3.10}$$

Note that the assignment $\mathrm{jac} : G(\mathbb{Q}) \to \mathcal{O}(\mathbb{H})$, which associates the function $z \mapsto \mathrm{jac}(\gamma, z)$ to an element $\gamma \in G(\mathbb{Q})$, descends naturally to the coset space $B(\mathbb{Z})\backslash G(\mathbb{Q})$, and even further to $B_u(\mathbb{Q})\backslash G(\mathbb{Q})$.

From (2.3.9) and (2.3.10) we see that $|\gamma \cdot z - \gamma \cdot z'| = |z - z'|$ just in the case that z and z' belong to the set $\{z \in \mathbb{H} \mid |\mathrm{jac}(\gamma, z)| = 1\}$ which we call the *isometric locus of* γ. The isometric locus of γ is a (Euclidean) semicircle just when $c(\gamma) > 0$, in which case its center is $\gamma^{-1} \cdot \infty$, and its radius is $\sqrt{\varrho(\gamma)}$, where

$$\varrho(\gamma)=\frac{P\det(\gamma)}{c(\gamma)^2}=\frac{ad-bc}{c^2} \tag{2.3.11}$$

in case $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ is a preferred representative for γ. We call $\varrho(\gamma)$ the *scaling factor* associated to γ.

2.4 Cosets

Given a subset $X\subset G(\mathbb{Q})$, we will write $\llbracket X\rrbracket$ for the set consisting of double cosets of the form $B(\mathbb{Z})\chi B(\mathbb{Z})$, for $\chi\in X$. We will write $\llbracket\chi\rrbracket$ as a shorthand for the particular double coset $B(\mathbb{Z})\chi B(\mathbb{Z})$.

$$\llbracket\chi\rrbracket=B(\mathbb{Z})\chi B(\mathbb{Z}),\quad \llbracket X\rrbracket=\{\llbracket\chi\rrbracket\in B(\mathbb{Z})\backslash G(\mathbb{Q})/B(\mathbb{Z})\mid\chi\in X\} \tag{2.4.1}$$

Similarly, we will write $\llbracket X$ for the set of right cosets of the form $B(\mathbb{Z})\chi$, for $\chi\in X$, and $\chi\mapsto\llbracket\chi$ will denote the natural map $X\to\llbracket X$, and the notations $X\rrbracket$ and $\chi\mapsto\chi\rrbracket$ will have the analogous meanings. Observe that $B(\mathbb{Z})$ is stable under the operations of inversion and conjugation (cf. (2.3.2)), and thus these operations descend naturally to the coset spaces $\llbracket G(\mathbb{Q})$ and $G(\mathbb{Q})\rrbracket$ and $\llbracket G(\mathbb{Q})\rrbracket$.

Recall from §2.3 that the assignment $\chi\mapsto\operatorname{jac}(\chi,z)$ descends to a well-defined association of holomorphic functions on $\mathbb{H}$ to right cosets of $B(\mathbb{Z})$ in $G(\mathbb{Q})$. Accordingly, we may safely set $\operatorname{jac}(\llbracket\chi\,,z)=\operatorname{jac}(\chi,z)$, when given some $\llbracket\chi\in\llbracket G(\mathbb{Q})$.

$$\begin{aligned}\operatorname{jac}(\cdot\,,z):\llbracket G(\mathbb{Q})&\to\mathcal{O}(\mathbb{H})\\ \llbracket\chi&\mapsto\operatorname{jac}(\llbracket\chi\,,z)\end{aligned} \tag{2.4.2}$$

Given $U\subset\llbracket G(\mathbb{Q})$ or $S\subset\llbracket G(\mathbb{Q})\rrbracket$ we set $U_\infty=U\cap\llbracket B(\mathbb{Q})$ and $U^\times=U-U_\infty$, and similarly, $S_\infty=S\cap\llbracket B(\mathbb{Q})\rrbracket$ and $S^\times=S-S_\infty$. In particular, $\llbracket G(\mathbb{Q})\rrbracket^\times$ denotes the set of double cosets $\llbracket\chi\rrbracket$, for $\chi\in G(\mathbb{Q})$, for which $c(\chi)\neq 0$.

Observe that the isometric loci (cf. §2.3) associated to $\chi,\chi'\in G(\mathbb{Q})$ coincide when $\llbracket\chi=\llbracket\chi'$. Thus we may regard these loci as naturally attached to right cosets of $B(\mathbb{Z})$ in $G(\mathbb{Q})$, and the locus attached to $\llbracket\chi$ is a Euclidean semicircle just when $\llbracket\chi\in\llbracket G(\mathbb{Q})^\times$. Also, the loci associated to cosets $\llbracket\chi\,,\llbracket\chi'\in\llbracket G(\mathbb{Q})^\times$ have the same radii when $\llbracket\chi\rrbracket=\llbracket\chi'\rrbracket$; that is to say, the function ϱ of (2.3.11) descends to a well-defined map $\llbracket G(\mathbb{Q})\rrbracket^\times\to\mathbb{Q}^+$. It is worth noting that the function $\chi\mapsto\varrho(\chi)$ actually satisfies an even stronger invariance condition.

Lemma 2.4.1. *The assignment $\llbracket\chi\rrbracket\mapsto\varrho\llbracket\chi\rrbracket$ descends naturally to a well-defined function on the double coset space $B_u(\mathbb{Q})\backslash G(\mathbb{Q})/B_u(\mathbb{Q})$. That is, we have $\varrho\llbracket T^\alpha\chi\rrbracket=\varrho\llbracket\chi T^\alpha\rrbracket=\varrho\llbracket\chi\rrbracket$ for any $\alpha\in\mathbb{Q}$ and $\chi\in G(\mathbb{Q})^\times$.*

The function ϱ is also invariant under inversion and conjugation, so that we have $\varrho\llbracket\chi\rrbracket = \varrho\llbracket\bar{\chi}\rrbracket = \varrho\left\llbracket\chi^{-1}\right\rrbracket$ for any $\chi \in G(\mathbb{Q})^{\times}$. For $\chi \in G(\mathbb{Q})^{\times}$ the identity (2.3.9) may be rewritten

$$\chi \cdot z - \chi \cdot z' = \frac{\varrho\llbracket\chi\rrbracket}{(z - \chi^{-1}\cdot\infty)(z' - \chi^{-1}\cdot\infty)}\,(z - z'), \tag{2.4.3}$$

and we may consider the limit as $z' \to \infty$, which yields a useful expression for $\chi \cdot z - \chi \cdot \infty$; viz.,

$$\chi \cdot z - \chi \cdot \infty = -\frac{\varrho\llbracket\chi\rrbracket}{z - \chi^{-1}\cdot\infty}. \tag{2.4.4}$$

Comparing with (2.3.10) and (2.4.4) we have

$$\mathrm{jac}(\llbracket\chi\,, z) = \frac{\varrho\llbracket\chi\rrbracket}{(z - \chi^{-1}\rrbracket\cdot\infty)^2} \tag{2.4.5}$$

for $\llbracket\chi \in \llbracket G(\mathbb{Q})^{\times}$ (and $\mathrm{jac}(\llbracket\chi\,, z) = 1$ otherwise).

The functions $c, d : G(\mathbb{Q}) \to \mathbb{N}$, of §2.3, descend to well-defined functions on the coset space $\llbracket G(\mathbb{Q}) = B(\mathbb{Z})\backslash G(\mathbb{Q})$. The function c descends further, to be well-defined on the double coset space $\llbracket G(\mathbb{Q})\rrbracket = B(\mathbb{Z})\backslash G(\mathbb{Q})/B(\mathbb{Z})$. For $K \in \mathbb{R}^+$ and U a subset of $\llbracket G(\mathbb{Q})$, define $U_{\leq K}$ to be the subset of U consisting of cosets $\llbracket\chi \in U$ for which $c(\chi) \leq K$ and $d(\chi) \leq K^2$. Analogously, for S a subset of $\llbracket G(\mathbb{Q})\rrbracket$, define $S_{\leq K}$ to be the subset of S consisting of double cosets $\llbracket\chi\rrbracket \in S$ for which $c(\chi) \leq K$.

$$U_{\leq K} = \left\{\llbracket\chi \in U \mid c(\chi) \leq K,\, d(\chi) \leq K^2\right\} \tag{2.4.6}$$

$$S_{\leq K} = \{\llbracket\chi\rrbracket \in S \mid c(\chi) \leq K\} \tag{2.4.7}$$

2.5 Groups

The group $G(\mathbb{Q})$ is the commensurator of $G(\mathbb{Z})$. Consequently, any group $\Gamma < G(\mathbb{R})$ that is commensurable with the modular group $G(\mathbb{Z})$ is automatically a subgroup of $G(\mathbb{Q})$, and is an example of a Fuchsian group of the first kind. For such a group Γ there is a natural way to equip the orbit space $\Gamma\backslash\mathbb{H}$ with the structure of a Riemann surface; we will denote this object by Y_Γ. Any group Γ commensurable with $G(\mathbb{Z})$ has parabolic elements, and any fixed point of any parabolic element of Γ lies on the (pointed) rational projective line $\hat{\mathbb{Q}} \subset \hat{\mathbb{R}}$. The orbit space $\mathsf{P}_\Gamma = \Gamma\backslash\hat{\mathbb{Q}}$ is the set of *cusps of* Γ. For Γ commensurable with $G(\mathbb{Z})$, the Riemann surface Y_Γ is not compact, but can be compactified in a natural way by adjoining a single point for each cusp of Γ. We will write X_Γ for the corresponding compact Riemann surface obtained by the adjunction of the cusps of Γ.

$$\mathsf{X}_\Gamma = \mathsf{Y}_\Gamma \cup \mathsf{P}_\Gamma = \Gamma\backslash\mathbb{H} \cup \Gamma\backslash\hat{\mathbb{Q}} \tag{2.5.1}$$

We will say that Γ is a *group of genus zero* in the case that X_Γ has genus zero as a Riemann surface. We set $\operatorname{genus}(\Gamma) = \operatorname{genus}(\mathsf{X}_\Gamma)$ for $\Gamma < G(\mathbb{R})$ commensurable with $G(\mathbb{Z})$.

We write $\hat{\mathbb{C}}$ for the *(pointed) complex projective line* $\mathbb{C} \cup \{\infty\}$, which we regard as a compact Riemann surface in the usual way. If Γ is a group for which the compact Riemann surface X_Γ has genus zero, then there is an isomorphism $\phi : \mathsf{X}_\Gamma \xrightarrow{\sim} \hat{\mathbb{C}}$ that witnesses this fact, and the set of such isomorphisms admits a simply transitive action by $G(\mathbb{C})$. Thus we may assume that ϕ maps the point of X_Γ corresponding to the *infinite cusp* $\Gamma \cdot \infty$ to the distinguished point $\infty \in \hat{\mathbb{C}}$. Such an isomorphism ϕ determines a Γ-invariant holomorphic function, f say, on $\mathbb{H}$, which admits an expression $f(z) = \sum_{n \geq -1} c(n/h)\,\mathrm{e}(nz/h)$ with $c(-1/h) \neq 0$ for some positive integer h called the *width of* Γ *at infinity*. After multiplying by $1/c(-1/h)$ we may assume that $c(-1/h) = 1$. A Γ-invariant function of the form $\mathrm{e}(-z/h) + \sum_{n \geq 0} c(n/h)\,\mathrm{e}(nz/h)$ which is holomorphic on $\mathbb{H}$ and induces an isomorphism $\mathsf{X}_\Gamma \xrightarrow{\sim} \hat{\mathbb{C}}$ will be called a *hauptmodul* for Γ. In case $c(0) = 0$ we say that f is a *normalized hauptmodul* for Γ.

The *Hecke congruence groups*, denoted $\Gamma_0(n)$ for $n \in \mathbb{Z}^+$, play a special rôle in our analysis.

$$\Gamma_0(n) = \left\{ \begin{bmatrix} a & b \\ cn & d \end{bmatrix} \mid a, b, c, d \in \mathbb{Z},\ ad - bcn = 1 \right\} \tag{2.5.2}$$

According to [CN79] the normalizer $N(\Gamma_0(n))$ of $\Gamma_0(n)$ in $G(\mathbb{R})$ is commensurable with $G(\mathbb{Z})$ and admits the description

$$N(\Gamma_0(n)) = \left\{ \begin{bmatrix} ae & b/h \\ cn/h & de \end{bmatrix} \mid a, b, c, d \in \mathbb{Z},\ e \in \mathbb{Z}^+,\ e \| n/h,\ ade - bcn/eh^2 = 1 \right\} \tag{2.5.3}$$

where $h = ((n, 24)$ is the largest divisor of 24 such that h^2 divides n (cf. §2.2). The expression (2.5.3) tells us, in particular, that $N(\Gamma_0(n))_\infty = \langle T^{1/h} \rangle$ for $h = ((n, 24)$.

Proposition 2.5.1. *Let $n \in \mathbb{Z}^+$. Then $N(\Gamma_0(n))$ acts transitively on $\hat{\mathbb{Q}}$. We have $N(\Gamma_0(n))_\infty = \Gamma_0(n)_\infty$ if and only if n is not divisible by 4 or 9.*

An important family of groups, each one commensurable with the modular group $G(\mathbb{Z})$, and each one containing and normalizing some $\Gamma_0(n)$, was introduced in [CN79]; these are the groups of *$n\|h$-type*, and we now recall their definition. For $n, h \in \mathbb{Z}^+$ with $h|n$, and for S a subgroup of $\operatorname{Ex}(n/h)$ (cf. §2.2), we define a group $\Gamma_0(n|h) + S$ by setting

$$\Gamma_0(n|h) + S = \left\{ \begin{bmatrix} ae & b/h \\ cn & de \end{bmatrix} \mid a, b, c, d \in \mathbb{Z},\ e \in S,\ ade - bcn/eh = 1 \right\}. \tag{2.5.4}$$

Evidently the group $\Gamma_0(n|h) + S$ contains $\Gamma_0(nh)$. In case $h|24$ it also normalizes $\Gamma_0(nh)$. Assume then that $h|24$. We say that $\Gamma_0(n|h) + S$ is of *Fricke type* if the element $\left[\begin{smallmatrix} 0 & -1 \\ nh & 0 \end{smallmatrix}\right]$ belongs to it.

The group $\Gamma_0(n\|h)+S$ is defined (cf. [CN79]) as the subgroup of $\Gamma_0(n|h)+S$ arising as the kernel of a certain morphism $\Gamma_0(n|h)+S\to\mathbb{Z}/h$ which factors through the canonical map $\Gamma_0(n|h)+S\to(\Gamma_0(n|h)+S)/\Gamma_0(nh)$. In order to describe it let S' denote the image of S in $\mathrm{Ex}(nh)$ under the natural injection $\mathrm{Ex}(n/h)\to\mathrm{Ex}(nh)$ (cf. §2.2). Then as generators for the quotient $(\Gamma_0(n|h)+S)/\Gamma_0(nh)$ we may take the cosets X, Y, and $W_{e'}$ for $e'\in S'$, given by

$$X=\begin{bmatrix}1&1/h\\0&1\end{bmatrix}\Gamma_0(nh),\quad Y=\begin{bmatrix}1&0\\n&1\end{bmatrix}\Gamma_0(nh),\quad W_{e'}=\begin{bmatrix}ae'&b\\cnh&de'\end{bmatrix}\Gamma_0(nh),\tag{2.5.5}$$

where the a, b, c and d in the definition of $W_{e'}$ are arbitrary integers for which $ade'-bcnh/e'=1$, and we may define $\Gamma_0(n\|h)+S$ to be the kernel of the composition

$$\Gamma_0(n|h)+S\to(\Gamma_0(n|h)+S)/\Gamma_0(nh)\xrightarrow{\lambda}\mathbb{Z}/h\tag{2.5.6}$$

where the map λ is determined by the requirements that $\lambda(W_{e'})=0$ for all $e'\in S'$, $\lambda(X)=1$, and $\lambda(Y)$ is 1 or -1 according as $\Gamma_0(n|h)+S$ is of Fricke type or not.

The group $\Gamma_0(n\|h)+S$ is also described in [Fer93], where it is denoted $(1/h)\Gamma_0(n|h)+e_1,e_2,\ldots$ for $S=\{1,e_1,e_2,\ldots\}$, although we should remark that certain variations on λ are considered there (in the case that h is even), so that the class of groups considered in [Fer93] properly contains the groups of $n\|h$-type.

2.6 Scalings

Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p}\in\mathsf{P}_\Gamma$ (cf. (2.5.1)) be a cusp of Γ. An element $p\in\mathsf{p}\subset\hat{\mathbb{Q}}$ will be called a *representative for* p.

Lemma 2.6.1. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $p\in\hat{\mathbb{Q}}$. Then there exists an element $\sigma_p\in G(\mathbb{Q})$ such that*

$$p=\sigma_p\cdot\infty,\quad(\sigma_p^{-1}\Gamma\sigma_p)_\infty=B(\mathbb{Z}).\tag{2.6.1}$$

Proof. In case $p=\infty$ the group Γ_∞ (cf. (2.3.7)) of elements in Γ that fix ∞ is infinite cyclic and generated by some element $\gamma_\infty\in B(\mathbb{Q})$. We claim that in fact γ_∞ lies in $B_u(\mathbb{Q})$, and thus equals T^α (cf. (2.3.4)) for some $\alpha\in\mathbb{Q}^+$. For otherwise $P\det(\gamma_\infty)=e$ for some $e>1$ (cf. (2.3.6)), so that $P\det(\gamma_\infty^n)=e^n$ for $n\in\mathbb{Z}$, since the restriction $P\det:B(\mathbb{Q})\to\mathbb{Z}^+$ is multiplicative. Then each power of γ_∞ lies in a distinct coset of the intersection $\Gamma\cap G(\mathbb{Z})$ in Γ since $P\det$ is invariant under multiplication by elements of $G(\mathbb{Z})$. This contradicts the hypothesis that Γ is commensurable with $G(\mathbb{Z})$, so we conclude that $\Gamma_\infty=\langle T^\alpha\rangle$ for some $\alpha\in\mathbb{Q}^+$. Then we may take $\sigma_\infty = [\alpha]$ (cf. (2.3.3)), for upon calculating $[\mu]T^\alpha[1/\mu] = T^{\mu\alpha}$ we find that

$$([1/\alpha]\Gamma[\alpha])_\infty = [1/\alpha]\Gamma_\infty[\alpha] = [1/\alpha]\langle T^\alpha\rangle[\alpha] = \langle T\rangle, \tag{2.6.2}$$

and the group $\langle T\rangle$ is just $B(\mathbb{Z})$, so the element $\sigma_\infty = [\alpha]$ satisfies the two conditions (2.6.1).

In case $p = a/c$ for coprime integers $a, c \in \mathbb{Z}$ with $c \neq 0$, we may choose $b, d \in \mathbb{Z}$ such that $ad - bc = 1$ and set $\sigma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Then $p = \sigma \cdot \infty$ and σ satisfies the first condition in (2.6.1). Now we rerun the argument of the paragraph above with $\sigma^{-1}\Gamma\sigma$ in place of Γ to find that $(\sigma^{-1}\Gamma\sigma)_\infty = \langle T^\alpha\rangle$ for some $\alpha \in \mathbb{Q}^+$, and $\sigma_p = \sigma[\alpha]$ satisfies the two conditions (2.6.1). This completes the proof of the lemma. □

An element $\sigma_p \in G(\mathbb{Q})$ satisfying the two conditions (2.6.1) of Lemma 2.6.1 will be called a *scaling element for* Γ *at the cusp representative* p. It is useful to have a replacement for the notion of scaling element that is independent of a choice of cusp representative.

Lemma 2.6.2. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$ *and let* $\mathsf{p} \in \mathsf{P}_\Gamma$ *be a cusp of* Γ*. Then there exists a coset* $\Sigma_\mathsf{p} \in \Gamma\backslash G(\mathbb{Q})$ *such that*

$$\mathsf{p} = \Sigma_\mathsf{p} \cdot \infty, \quad (\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{p})_\infty = B(\mathbb{Z}). \tag{2.6.3}$$

Proof. Let $p \in \mathsf{p} \subset \hat{\mathbb{Q}}$ be a representative for p and let $\sigma_p \in G(\mathbb{Q})$ satisfy the conditions (2.6.1) of Lemma 2.6.1. Then the coset $\Sigma_\mathsf{p} = \Gamma\sigma_p$ satisfies the conditions (2.6.3). □

A coset $\Sigma_\mathsf{p} \in \Gamma\backslash G(\mathbb{Q})$ satisfying the two conditions (2.6.3) of Lemma 2.6.2 will be called a *scaling coset for* Γ *at the cusp* p. Any element $\sigma \in \Sigma_\mathsf{p}$ is a scaling element for Γ at the representative $\sigma\cdot\infty$ for the cusp p, and if σ_p is a scaling element for Γ at the cusp representative p then the coset $\Gamma\sigma_p$ is a scaling coset for Γ at the cusp $\mathsf{p} = \Gamma \cdot p$ represented by p.

Scaling cosets are determined only up to right multiplication by elements of $B_u(\mathbb{Q})$. Indeed, both the conditions defining the notion of scaling coset are invariant under the replacement of Σ_p by $\Sigma_\mathsf{p}T^\alpha$ for some $\alpha \in \mathbb{Q}$ (cf. (2.3.4)). A set $\{\Sigma_\mathsf{p} \mid \mathsf{p} \in \mathsf{P}_\Gamma\} \subset \Gamma\backslash G(\mathbb{Q})$ such that Σ_p is a scaling coset for Γ at p for each $\mathsf{p} \in \mathsf{P}_\Gamma$ will be called a *system of scaling cosets for* Γ.

The problem of constructing scaling cosets for a group Γ may be viewed in the following way. The group $G(\mathbb{Q})$ acts transitively on $\hat{\mathbb{Q}}$, and the subgroup $B(\mathbb{Q})$ is just the stabilizer of the distinguished point $\infty \in \hat{\mathbb{Q}}$. Thus the map $\Gamma\backslash G(\mathbb{Q}) \to \mathsf{P}_\Gamma$ given by $\Gamma\sigma \mapsto \Gamma\sigma\cdot\infty$ induces an isomorphism $\Gamma\backslash G(\mathbb{Q})/B(\mathbb{Q}) \cong \mathsf{P}_\Gamma$ which we may regard as identifying the sets $\Gamma\backslash G(\mathbb{Q})/B(\mathbb{Q})$ and P_Γ. Now $B(\mathbb{Q})$ is naturally isomorphic to the semidirect product $B_u(\mathbb{Q}) \rtimes B_d(\mathbb{Q})$ (cf. §2.5), so the natural map $\Gamma\backslash G(\mathbb{Q}) \to \Gamma\backslash G(\mathbb{Q})/B(\mathbb{Q})$ factors through $\Gamma\backslash G(\mathbb{Q})/B_u(\mathbb{Q})$, yielding a naturally defined sequence

$$\Gamma\backslash G(\mathbb{Q}) \to \Gamma\backslash G(\mathbb{Q})/B_u(\mathbb{Q}) \to \Gamma\backslash G(\mathbb{Q})/B(\mathbb{Q}) \cong \mathsf{P}_\Gamma \tag{2.6.4}$$

where the fibres of the second map are torsors for the diagonal group $B_d(\mathbb{Q}) \simeq \mathbb{Q}^+$. Given a cusp $\mathsf{p} \in \mathsf{P}_\Gamma$, any preimage of p in $\Gamma\backslash G(\mathbb{Q})$ under the composition (2.6.4) is a coset of Γ satisfying the first condition defining a scaling coset for Γ at p. In order to satisfy also the second condition we should multiply this coset by $[\mu]$ (cf. (2.3.3)) for some (uniquely defined) $\mu \in \mathbb{Q}^+$. Thus the two conditions defining scaling cosets define a section of the $B_d(\mathbb{Q})$-bundle $\Gamma\backslash G(\mathbb{Q})/B_u(\mathbb{Q}) \to \mathsf{P}_\Gamma$. We denote this map $\mathsf{p} \mapsto \mathfrak{S}_\mathsf{p}$. We may naturally identify $\mathfrak{S}_\mathsf{p}$ with the set of scaling cosets for Γ at p, for these are precisely the preimages of $\mathfrak{S}_\mathsf{p}$ in $\Gamma\backslash G(\mathbb{Q})$ under the first map of (2.6.4).

$$\begin{aligned} \mathsf{P}_\Gamma &\to \Gamma\backslash G(\mathbb{Q})/B_u(\mathbb{Q}) \\ \mathsf{p} &\mapsto \mathfrak{S}_\mathsf{p} = \{\Sigma_\mathsf{p} T^\alpha \mid \alpha \in \mathbb{Q}\} \end{aligned} \tag{2.6.5}$$

One says that Γ has *width one at infinity* in the case that $\Gamma_\infty = B(\mathbb{Z})$. Observe that Γ has width one at infinity if and only if Γ is a scaling coset for itself at the infinite cusp $\Gamma \cdot \infty$; that is, if and only if $\Gamma \in \mathfrak{S}_{\Gamma\cdot\infty}$. If Γ does not have width one at infinity then there is a unique $\mu \in \mathbb{Q}^+$ with the property that $\Gamma[\mu]$ is a scaling coset for Γ at the infinite cusp $\Gamma \cdot \infty$, and then the group $\Gamma^{[\mu]} = [1/\mu]\Gamma[\mu]$ is a group with width one at infinity.

When engaged in the task of computing Fourier coefficients of modular forms for a group Γ say, one frequently has use for double coset spaces of the form $\Gamma_p\backslash\Gamma/\Gamma_q$, for some $p, q \in \hat{\mathbb{Q}}$, where Γ_p denotes the stabilizer in Γ of p. If σ_p and σ_q are scaling elements for Γ at p and q, respectively, then we have $\sigma_p^{-1}\Gamma_p\sigma_p = B(\mathbb{Z})$, and similarly with q in place of p, so the set of *translates*

$$\sigma_p^{-1}(\Gamma_p\backslash\Gamma/\Gamma_q)\sigma_q = \{\sigma_p^{-1}\Gamma_p\gamma\Gamma_q\sigma_q \mid \gamma \in \Gamma\} \tag{2.6.6}$$

is in fact a set of double cosets of $B(\mathbb{Z})$. In the notation of §2.4 we have

$$\sigma_p^{-1}(\Gamma_p\backslash\Gamma/\Gamma_q)\sigma_q = [\![\sigma_p^{-1}\Gamma\sigma_q]\!] = \left\{ [\![\sigma_p^{-1}\gamma\sigma_q]\!] \mid \gamma \in \Gamma \right\}. \tag{2.6.7}$$

Set $\mathsf{p} = \Gamma \cdot p$ and $\mathsf{q} = \Gamma \cdot q$, and set $\Sigma_\mathsf{p} = \Gamma\sigma_p$ and $\Sigma_\mathsf{q} = \Gamma\sigma_q$, so that Σ_p and Σ_q are scaling cosets for Γ at p and q, respectively. Then we have $\sigma_p^{-1}\Gamma\sigma_q = \Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}$, so that these translates depend only on cusps, and not on cusp representatives. We see from this discussion that for any pair of cusps $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$, with scaling cosets Σ_p and Σ_q, respectively, the set $\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}$ is a union of double cosets of $B(\mathbb{Z})$. The assignment $(\mathsf{p}, \mathsf{q}) \mapsto \Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}$ is sensitive to the choice of scaling cosets Σ_p and Σ_q, to the extent that a different choice will replace $\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}$ with a set of the form $T^\alpha\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}T^\beta$ for some $\alpha, \beta \in \mathbb{Q}$.

In general there may be no canonical choice of scaling coset for a particular group at a particular cusp, but there are situations in which some choices might be preferred over others. For example, it is natural to take $\Sigma_{\Gamma\cdot\infty} = \Gamma$ in case Γ has width one at infinity, for we then have $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} = \Sigma_{\mathsf{p}}^{-1}$ when $\mathsf{q} = \Gamma\cdot\infty$, and $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} = \Sigma_{\mathsf{q}}$ when $\mathsf{p} = \Gamma\cdot\infty$, and $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} = \Gamma$ when $\mathsf{p} = \mathsf{q} = \Gamma\cdot\infty$. We shall always take the scaling coset $\Sigma_{\Gamma\cdot\infty}$, for Γ at the infinite cusp, to be of the form $\Gamma[\mu]$ for $\mu \in \mathbb{Q}^+$. This value μ is uniquely determined.

Observe that if Σ_{p} and Σ_{q} are scaling cosets for Γ at p and q, respectively, then the intersection $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} \cap B(\mathbb{Q}) = (\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}})_\infty$ can be non-empty only in the case that $\mathsf{p} = \mathsf{q}$. For p and q cusps for Γ we define $\delta_{\Gamma|\mathsf{p},\mathsf{q}}$ to be 1 or 0 according as $\mathsf{p} = \mathsf{q}$ or not.

Given a system $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ of scaling cosets for Γ, we write Γ^{p} as a shorthand for the group obtained as the conjugate of Γ by Σ_{p}.

$$\Gamma^{\mathsf{p}} = \Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}} \tag{2.6.8}$$

The notation (2.6.8) suppresses the dependence on the choice of scaling coset: a different choice will replace Γ^{p} with a group of the form $T^{-\alpha}\Gamma^{\mathsf{p}}T^{\alpha}$ for some $\alpha \in \mathbb{Q}$.

2.7 Integrals

Write $\mathcal{O}(\mathbb{H})$ for the ring of holomorphic functions on the upper half plane $\mathbb{H}$.

For $\kappa \in \mathbb{Z}$, we define the *weight 2κ (right) action of $G(\mathbb{Q})$ on $\mathcal{O}(\mathbb{H})$*, to be denoted $(f,\chi) \mapsto f|^{\kappa}\chi$, by setting

$$\left(f|^{\kappa}\chi\right)(z) = f(\chi\cdot z)\,\mathrm{jac}(\chi,z)^{\kappa} \tag{2.7.1}$$

for $f \in \mathcal{O}(\mathbb{H})$ and $\chi \in G(\mathbb{Q})$. For Γ a group commensurable with $G(\mathbb{Z})$ (and hence a subgroup of $G(\mathbb{Q})$), we call $f \in \mathcal{O}(\mathbb{H})$ an *unrestricted modular form of weight 2κ for* Γ in case it is a fixed point for the weight 2κ action of Γ. Suppose f is an unrestricted modular form of weight 2κ for Γ. Then for $X \in \Gamma\backslash G(\mathbb{Q})$ a right coset of Γ in $G(\mathbb{Q})$ we may define a function $(f\|^{\kappa}_{\Gamma}X)(z)$ by setting

$$f\|^{\kappa}_{\Gamma}X = f|^{\kappa}\chi \tag{2.7.2}$$

where χ is any representative for the coset $X \in \Gamma\backslash G(\mathbb{Q})$. Let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ and let $\Sigma_{\mathsf{p}} \in \Gamma\backslash G(\mathbb{Q})$ a scaling coset for Γ at p. Then for f an unrestricted modular form of weight 2κ for Γ we define $f_{|\mathsf{p}} \in \mathcal{O}(\mathbb{H})$ by setting $f_{|\mathsf{p}} = f\|^{\kappa}_{\Gamma}\Sigma_{\mathsf{p}}$. Then $f_{|\mathsf{p}}(z+1) = f_{|\mathsf{p}}(z)$ for all $z \in \mathbb{H}$, by the defining properties of Σ_{p} (cf. §2.6), so we have

$$f_{|\mathsf{p}}(z)=\sum_{n\in\mathbb{Z}}c_{|\mathsf{p}}(n)\,\mathrm{e}(nz) \tag{2.7.3}$$

for some $c_{|\mathsf{p}}(n)\in\mathbb{Z}$. We call the right hand side of (2.7.3) the *Fourier expansion of* f *at* p *with respect to* Σ_{p}, and we say that f is *meromorphic at* p if the right hand side of (2.7.3) is a meromorphic function of $q=\mathrm{e}(z)$ in a neighborhood of $q=0$. A different choice of scaling coset will replace $c_{|\mathsf{p}}(n)$ with $\mathrm{e}(n\alpha)c_{|\mathsf{p}}(n)$ for some $\alpha\in\mathbb{Q}$, so the notion of being meromorphic at p is independent of the choice of scaling coset at p.

We define $M_\kappa(\Gamma)$ to be the vector space consisting of unrestricted modular forms of weight 2κ for Γ that are meromorphic at the cusps of Γ. In a slight departure from standard convention we call $M_\kappa(\Gamma)$ the space of *modular forms of weight* 2κ *for* Γ. We define $S_\kappa(\Gamma)$ to be the vector space consisting of modular forms of weight 2κ for Γ that vanish at the cusps of Γ. We call $S_\kappa(\Gamma)$ the space of *cusp forms of weight* 2κ *for* Γ. The vector space $S_\kappa(\Gamma)$ is the zero vector space when $\kappa\leq 0$.

Let $\kappa\in\mathbb{Z}$ such that $\kappa\leq 0$, and let $w\in\mathbb{H}\cup\hat{\mathbb{Q}}$. Given a holomorphic function $g(z)$ on $\mathbb{H}$ with sufficiently rapid decay as the imaginary part of z tends to ∞ we define a function $I^\kappa_w g$ on $\mathbb{C}$ by setting $I^\kappa_\infty g$ to be the zero function and by setting

$$(I^\kappa_w g)(z)=2\pi\mathbf{i}\int_w^\infty g(\xi)(2\pi\mathbf{i}(\xi-z))^{(-2\kappa)}\mathrm{d}\xi \tag{2.7.4}$$

for $w\neq\infty$, where the integral is taken over the vertical line $\{w+\mathbf{i}t\mid t\in\mathbb{R}^+\}$. In our applications g will either be a cusp form (for some $\Gamma<PSL_2(\mathbb{R})$ commensurable with $PSL_2(\mathbb{Z})$) or will admit a finite power series expansion of the form $g(z)=\sum_{n=1}^N c(n)\,\mathrm{e}(nz)$. In these cases the convergence of the integral defining $I^\kappa_w g$ is clear.

We define a closely related operator $g\mapsto J^\kappa_w g$ by setting

$$(J^\kappa_w g)(z)=\overline{(I^\kappa_w g)(\bar{z})}. \tag{2.7.5}$$

Observe that when $\kappa\leq 0$, the functions $(I^\kappa_w g)(z)$ and $(J^\kappa_w g)(z)$ are polynomials of degree at most -2κ in z . For $\chi\in G(\mathbb{Q})$ we have

$$(I^\kappa_w g)|^\kappa\chi=I^\kappa_{w\cdot\chi}\left(g|^{1-\kappa}\chi\right)-I^\kappa_{\infty\cdot\chi}\left(g|^{1-\kappa}\chi\right) \tag{2.7.6}$$

where $w\cdot\chi=\chi^{-1}\cdot w$ denotes the natural right action of $G(\mathbb{Q})$ on $\mathbb{H}\cup\hat{\mathbb{Q}}$. In particular, if $g\in M_{1-\kappa}(\Gamma)$ for some group Γ then we have $(I^\kappa_w g)|^\kappa\gamma=I^\kappa_{w\cdot\gamma}g-I^\kappa_{\infty\cdot\gamma}g$ for $\gamma\in\Gamma$.

Lemma 2.7.1. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$*, let* $\kappa\in\mathbb{Z}$ *such that* $\kappa\leq 0$*, and let* $g\in S_{1-\kappa}(\Gamma)$*. Then we have* $I^\kappa_p g=0$ *for all* $p\in\hat{\mathbb{Q}}$ *if and only if* $g=0$*.*

Proof. Certainly $I_p^\kappa g = 0$ for all $p \in \hat{\mathbb{Q}}$ if $g = 0$, so let $g \in S_{1-\kappa}(\Gamma)$ such that $I_p^\kappa g = 0$ for all $p \in \hat{\mathbb{Q}}$. Following the proof of Lemma 3.2 in [Nie74] we define a function $f(z)$ by setting

$$f(z) = (I_z^\kappa g)(z) = 2\pi\mathbf{i}\int_z^\infty g(\xi)(2\pi\mathbf{i}(\xi - z))^{(-2\kappa)}\mathrm{d}\xi. \tag{2.7.7}$$

The identity (2.7.6) implies $f|^\kappa\gamma = f - I^\kappa_{\infty\cdot\gamma}g$ for $\gamma \in \Gamma$, so that f is an unrestricted modular form of weight 2κ for Γ. We should consider the behavior of f at the cusps of Γ. Let $\mathsf{p} \in \mathsf{P}_\Gamma$, and let Σ_p be a scaling coset for Γ at p. Then we have

$$g_{|\mathsf{p}}(\xi) = \sum_{n>0} b_{|\mathsf{p}}(n)\,\mathrm{e}(n\xi) \tag{2.7.8}$$

for some $b_{|\mathsf{p}}(n) \in \mathbb{C}$, where $g_{|\mathsf{p}} = g\|_\Gamma^{1-\kappa}\Sigma_\mathsf{p}$ (cf. (2.7.2)). In order to analyze $f_{|\mathsf{p}} = f\|_\Gamma^\kappa\Sigma_\mathsf{p}$ we choose $\sigma \in \Sigma_\mathsf{p}$ and compute

$$f|^\kappa\sigma = I_z^\kappa(g_{|\mathsf{p}}) - I^\kappa_{\infty\cdot\sigma}(g_{|\mathsf{p}}) \tag{2.7.9}$$

by (2.7.6). Now the term $I^\kappa_{\infty\cdot\sigma}(g_{|\mathsf{p}})$ vanishes, since we have

$$\begin{aligned}(I^\kappa_{\infty\cdot\sigma}(g_{|\mathsf{p}}))(z) &= 2\pi\mathbf{i}^{1-2\kappa}\int_{\sigma^{-1}\cdot\infty}^\infty g(\sigma\cdot\xi)\,\mathrm{jac}(\sigma,\xi)^{1-\kappa}(\xi - z)^{(-2\kappa)}\mathrm{d}\xi \\ &= 2\pi\mathbf{i}^{1-2\kappa}\int_{\sigma^{-1}\cdot\infty}^\infty g(\sigma\cdot\xi)(\sigma\cdot\xi - \sigma\cdot z)^{(-2\kappa)}\,\mathrm{jac}(\sigma,z)^\kappa\mathrm{d}(\sigma\cdot\xi) \\ &= -(I^\kappa_{\sigma\cdot\infty}g)\,\mathrm{jac}(\sigma,z)^\kappa,\end{aligned} \tag{2.7.10}$$

which vanishes by our hypothesis on g. So we have the Fourier expansion

$$f_{|\mathsf{p}}(z) = \sum_{n>0} c_{|\mathsf{p}}(n)\,\mathrm{e}(nz) \tag{2.7.11}$$

for f at p, where $c_{|\mathsf{p}}(n)$ is determined by $b_{|\mathsf{p}}(n)$ according to the formula

$$c_{|\mathsf{p}}(n) = b_{|\mathsf{p}}(n)2\pi\mathbf{i}(I_0^\kappa g)(0) = -n^{2\kappa-1}b_\mathsf{p}(n). \tag{2.7.12}$$

We conclude that f is a cusp form of weight κ. Since $\kappa \leq 0$ we have $S_\kappa(\Gamma) = 0$, and so $c_{|\mathsf{p}}(n) = 0$ for all $n \in \mathbb{Z}$ and all $\mathsf{p} \in \mathsf{P}_\Gamma$, and so $b_{|\mathsf{p}}(n) = 0$ for all $n \in \mathbb{Z}$ and all $\mathsf{p} \in \mathsf{P}_\Gamma$. That is to say, g vanishes identically, which is what we required to show. □

Let $\kappa \in \mathbb{Z}$ such that $\kappa \leq 0$. Following Niebur (cf. [Nie74]) we say that a holomorphic function $f \in \mathcal{O}(\mathbb{H})$ is an *unrestricted automorphic integral of weight* 2κ *for* Γ in case there exists a cusp form $g \in S_{1-\kappa}(\Gamma)$ such that

$$f|^{\kappa}\gamma = f - J^{\kappa}_{\infty\cdot\gamma}g \tag{2.7.13}$$

for each $\gamma \in \Gamma$. According to Lemma 2.7.1 the cusp form $g \in S_{1-\kappa}(\Gamma)$ satisfying (2.7.13) is uniquely determined. We call it the *cusp form associated to the unrestricted automorphic integral* f. The operator $f \mapsto f\|^{\kappa}_{\Gamma}X$, for X a right coset of Γ, may be extended to unrestricted automorphic integrals as follows. Suppose f is an unrestricted automorphic integral of weight 2κ for Γ with associated cusp form g. Then for $X \in \Gamma\backslash G(\mathbb{Q})$ we may define $f\|^{\kappa}_{\Gamma}X \in \mathcal{O}(\mathbb{H})$ by setting

$$f\|^{\kappa}_{\Gamma}X = (f - J^{\kappa}_{\chi\cdot\infty}g)|^{\kappa}\chi \tag{2.7.14}$$

where χ is any representative for the coset $X \in \Gamma\backslash G(\mathbb{Q})$. The transformation properties (2.7.6) and (2.7.13) confirm that the function $f\|^{\kappa}_{\Gamma}X$ is independent of the choice of coset representative χ. Let Σ_{p} be a scaling coset for Γ at a cusp $\mathsf{p} \in \mathsf{P}_{\Gamma}$, and define $f_{|\mathsf{p}}$ by setting

$$f_{|\mathsf{p}} = f\|^{\kappa}_{\Gamma}\Sigma_{\mathsf{p}}. \tag{2.7.15}$$

By the defining properties of scaling cosets we find that $f_{|\mathsf{p}}(z+1) = f_{|\mathsf{p}}(z)$ for all $z \in \mathbb{H}$, and so we have

$$f_{|\mathsf{p}}(z) = \sum_{n\in\mathbb{Z}} c_{|\mathsf{p}}(n)\,\mathrm{e}(nz) \tag{2.7.16}$$

for some $c_{|\mathsf{p}}(n) \in \mathbb{C}$. We call (2.7.16) the *Fourier expansion of* f *at* p *with respect to* Σ_{p}, and we say that f is *meromorphic at* p if the right hand side of (2.7.16) is a meromorphic function of $q = \mathrm{e}(z)$ in a neighborhood of $q = 0$.

We define $I_{\kappa}(\Gamma)$ to be the vector space consisting of unrestricted automorphic integrals of weight 2κ for Γ that are meromorphic at every cusp of Γ. We call $I_{\kappa}(\Gamma)$ the space of *automorphic integrals of weight* 2κ *for* Γ.

The space $M_{\kappa}(\Gamma)$ of modular forms of weight 2κ for Γ is a subspace of $I_{\kappa}(\Gamma)$ by definition. Lemma 2.7.1 may be regarded as stating that $M_{\kappa}(\Gamma)$ is the kernel of the map $I_{\kappa}(\Gamma) \to S_{1-\kappa}(\Gamma)$ which sends an automorphic integral f to its associated cusp form g (cf. (2.7.13)). We will see in §4.5 that the map $I_{\kappa}(\Gamma) \to S_{1-\kappa}(\Gamma)$ is surjective, so that we have an exact sequence of vector spaces

$$0 \to M_{\kappa}(\Gamma) \to I_{\kappa}(\Gamma) \to S_{1-\kappa}(\Gamma) \to 0. \tag{2.7.17}$$

The sequence (2.7.17) is trivially exact in case $\kappa > 0$, for then $S_{1-\kappa}(\Gamma) = 0$, so that an automorphic integral of weight 2κ is automatically a modular form. Indeed, and more generally, if $\kappa \in \mathbb{Z}$ is chosen so that $S_{1-\kappa}(\Gamma) = 0$, then the spaces $M_\kappa(\Gamma)$ and $I_\kappa(\Gamma)$ coincide by the definition (cf. (2.7.13)) of an automorphic integral.

3 Rademacher sums

In this section we associate Rademacher sums, and conjugate Rademacher sums (cf. §1.3), of arbitrary even integer weight to triples $(\Gamma, \mathsf{p}, \mathsf{q})$ where Γ is a group commensurable with the modular group and p and q are cusps for Γ. We derive explicit expressions for the Fourier expansions of these functions, and study how they transform under the action of Γ.

3.1 Construction

Given $\kappa, m \in \mathbb{Z}$ and $\llbracket \chi \in \llbracket G(\mathbb{Q})$, define a function $z \mapsto P^{\kappa(m)}_{\llbracket \chi}(z)$ on $\mathbb{H}$ by setting

$$P^{\kappa(m)}_{\llbracket \chi}(z) = \mathrm{e}(m \llbracket \chi \cdot z)\, \mathrm{jac}(\llbracket \chi\, , z)^{\kappa}. \tag{3.1.1}$$

Given $U \subset \llbracket G(\mathbb{Q})$, define the *holomorphic Poincaré series of weight 2κ and order m associated to U*, to be denoted $P^{\kappa(m)}_{U}(z)$, by setting

$$P^{\kappa(m)}_{U}(z) = \lim_{K\to\infty} \sum_{\llbracket \chi \in U_{\leq K}} P^{\kappa(m)}_{\llbracket \chi}(z) \tag{3.1.2}$$

where $U_{\leq K}$ is as in (2.4.6). Of course, the limit defining $P^{\kappa(m)}_{U}(z)$ may or may not converge depending on the choice of κ, m, and U. If $\kappa > 1$ and $m \geq 0$ then for $U = \llbracket \Gamma$ say, for Γ a group commensurable with $G(\mathbb{Z})$, the limiting sum in (3.1.2) converges absolutely and uniformly on compact subsets of $\mathbb{H}$ (cf. Lemma 3.4.2) and the symbols $\lim_{K\to\infty}$ and ${}_{\leq K}$ are unnecessary. If $\kappa \leq 0$ then even with the limit the expression on the right hand side of (3.1.2) fails to converge, and the Poincaré series $P^{\kappa(m)}_{U}(z)$ is not defined.

As was originally demonstrated by Rademacher in [Rad39], the non-convergence of the Poincaré series $P^{\kappa(m)}_{U}(z)$ when $\kappa = 0$ can be circumvented by replacing $P^{0(m)}_{\llbracket \chi}(z)$ with $P^{0(m)}_{\llbracket \chi}(z) - P^{0(m)}_{0,\llbracket \chi}$ in the right-hand side of (3.1.2) where

$$P^{0(m)}_{\llbracket \chi}(z) - P^{0(m)}_{0,\llbracket \chi} = \mathrm{e}(m \llbracket \chi \cdot z) - \mathrm{e}(m \llbracket \chi \cdot \infty) \tag{3.1.3}$$

in case $\llbracket \chi \in \llbracket G(\mathbb{Q})^{\times}$, and $P^{0(m)}_{0,\llbracket \chi} = 0$ otherwise. Rademacher showed for $m = -1$ (cf. loc. cit.) that the resulting expression is a (conditionally) convergent series which recovers a function invariant for the modular group $G(\mathbb{Z})$ in the case that $U = \llbracket G(\mathbb{Z})$. Generalizations of

Rademacher's construction were given by Knopp (cf. [Kno62a]) and Niebur (cf. [Nie74]) so as to obtain conditionally convergent Poincaré series for arbitrary real non-positive weights, for various subgroups of $G(\mathbb{R}) = PSL_2(\mathbb{R})$.

With the constructions of Rademacher, Knopp and Niebur in mind, we introduce the *Rademacher component function of weight* 2κ *and order* m *associated to* $\llbracket\chi$, denoted $z \mapsto R^{\kappa(m)}_{\llbracket\chi}(z)$ and defined for $\kappa, m \in \mathbb{Z}$ and $\llbracket\chi \in \llbracket G(\mathbb{Q})$, by setting

$$R^{\kappa(m)}_{\llbracket\chi}(z) = \mathrm{e}(m\llbracket\chi\cdot z)\,\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z)\,\mathrm{jac}(\llbracket\chi\,,z)^{\kappa}, \tag{3.1.4}$$

where $\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z)$ is the *Rademacher regularization factor of weight* 2κ, given by

$$\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z) = \Phi(1-2\kappa, 2-2\kappa, m\llbracket\chi\cdot\infty - m\llbracket\chi\cdot z)(2\pi\mathbf{i}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty))^{1-2\kappa} \tag{3.1.5}$$

in case $\chi \in G(\mathbb{Q})^{\times}$, and $\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z) = 1$ otherwise (cf. (2.1.3)). The function $R^{\kappa(m)}_{\llbracket\chi}(z)$ is entire in case $\chi \in B(\mathbb{Q})$, and is otherwise holomorphic away from $z = \chi^{-1}\cdot\infty$. For $U \subset \llbracket G(\mathbb{Q})$ we define the *classical Rademacher sum of weight* 2κ *and order* m *associated to* U, to be denoted $R^{\kappa(m)}_{U}(z)$, by setting

$$R^{\kappa(m)}_{U}(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} R^{\kappa(m)}_{\llbracket\chi}(z). \tag{3.1.6}$$

Again, the expression defining $R^{\kappa(m)}_{U}(z)$ will not converge for all choices of κ, m and U. Employing the Kummer transformation (2.1.4) we may rewrite the Rademacher component function $R^{\kappa(m)}_{\llbracket\chi}(z)$ as

$$R^{\kappa(m)}_{\llbracket\chi}(z) = \mathrm{e}(m\llbracket\chi\cdot\infty)\,\mathrm{e}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty, 1-2\kappa)\,\mathrm{jac}(\llbracket\chi\,,z)^{\kappa} \tag{3.1.7}$$

in case $\llbracket\chi \in \llbracket G(\mathbb{Q})^{\times}$, where $\mathrm{e}(z,s)$ is the generalized exponential function of (2.1.6). Since we restrict attention to the case that κ is an integer (cf. [Nie74] for non-integral weights), the generalized exponential in (3.1.7) may in turn be written in terms of the partial exponential function $\mathrm{e}(z)_{<K}$ (cf. §2.1). We thus obtain the identity

$$\mathrm{e}(m\llbracket\chi\cdot z)\,\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z) = \mathrm{e}(m\llbracket\chi\cdot z) - \mathrm{e}(m\llbracket\chi\cdot\infty)\,\mathrm{e}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty)_{<1-2\kappa} \tag{3.1.8}$$

when $\chi \in G(\mathbb{Q})^{\times}$. In particular, we have $\mathrm{r}^{\kappa}(m,\llbracket\chi\,,z) = 1$ when κ is positive, so that the Rademacher sum $R^{\kappa(m)}_{U}(z)$ is exactly (albeit formally) the Poincaré series $P^{\kappa(m)}_{U}(z)$ when $\kappa > 0$.

The case that $\kappa = 0$ is special, for the corresponding Poincaré series, if it were to converge, would define a Γ-invariant function on $\mathbb{H}$ for any group Γ satisfying $U\Gamma = U$. With $\kappa = 0$ we have $\mathrm{r}^0(m, \llbracket\chi\,, z) = 1 - \mathrm{e}(m \llbracket\chi \cdot \infty - m \llbracket\chi \cdot z)$ when $\chi \in G(\mathbb{Q})^\times$, and hence

$$R^{0(m)}_{\llbracket\chi}(z) = P^{0(m)}_{\llbracket\chi}(z) - P^{0(m)}_{0,\llbracket\chi} \tag{3.1.9}$$

for all $\llbracket\chi \in \llbracket G(\mathbb{Q})^\times$. In other words, the Rademacher sum $R^{0(m)}_U(z)$ is given by

$$\begin{aligned} R^{0(m)}_U(z) &= \lim_{K\to\infty} \sum_{\llbracket\chi \in U^\times_{\leq K}} P^{0(m)}_{\llbracket\chi}(z) - P^{0(m)}_{0,\llbracket\chi} \\ &= \sum_{\llbracket\chi \in U_\infty} \mathrm{e}(m \llbracket\chi \cdot z) + \lim_{K\to\infty} \sum_{\llbracket\chi \in U^\times_{\leq K}} \mathrm{e}(m \llbracket\chi \cdot z) - \mathrm{e}(m \llbracket\chi \cdot \infty), \end{aligned} \tag{3.1.10}$$

with the latter expression holding at least when there are only finitely many cosets of $B(\mathbb{Z})$ in U_∞. This expression (3.1.10), with $m = -1$ and $U = \llbracket G(\mathbb{Z})$, is the series originally considered by Rademacher in [Rad39].

It is interesting to consider the function obtained by substituting $\bar{z}$ for z in the right hand side of (3.1.6). We define the *conjugate Rademacher sum of weight 2κ and order m associated to U*, to be denoted $\bar{R}^{\kappa(m)}_U(z)$, by setting

$$\bar{R}^{\kappa(m)}_U(z) = \lim_{K\to\infty} \sum_{\llbracket\chi \in U_{\leq K}} \bar{R}^{\kappa(m)}_{\llbracket\chi}(z), \tag{3.1.11}$$

where $\bar{R}^{\kappa(m)}_{\llbracket\chi}(z) = R^{\kappa(m)}_{\llbracket\chi}(\bar{z})$ by definition. Evidently, the conjugate Rademacher sum $\bar{R}^{\kappa(m)}_U(z)$ is an anti-holomorphic function on $\mathbb{H}$ whenever U, κ and m are such that the limit in (3.1.11) converges locally uniformly in z. At first glance it appears that we should recover the classical Rademacher sum $R^{\kappa(m)}_U(z)$ from the conjugate Rademacher sum according to the identity $R^{\kappa(m)}_U(z) = \bar{R}^{\kappa(m)}_U(\bar{z})$, but as we shall see in §3.4 the relationship between the holomorphic functions $R^{\kappa(m)}_U(z)$ and $\bar{R}^{\kappa(m)}_U(\bar{z})$ is generally more interesting that this.

We typically take U to be of the form $U = \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ where $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ is a system of scaling cosets (cf. §2.6) for some group Γ commensurable with the modular group, and $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ are cusps of Γ. In this case we write $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for $R^{\kappa(m)}_U(z)$, and similarly for the holomorphic Poincaré series, suppressing the choice of scaling cosets from notation. A change in the choice of Σ_{p} and Σ_{q} replaces $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ with a function of the form $\mathrm{e}(\alpha)R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z+\beta)$ for some $\alpha, \beta \in \mathbb{Q}$, and similarly for the holomorphic Poincaré series $P^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$.

In the case that p or q is the *infinite cusp* $\Gamma \cdot \infty$ we omit it from notation, writing $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ for $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\Gamma\cdot\infty}(z)$, and $R^{\kappa(m)}_{\Gamma|\mathsf{q}}(z)$ for $R^{\kappa(m)}_{\Gamma,\Gamma\cdot\infty|\mathsf{q}}(z)$, and similarly for the holomorphic Poincaré series. The functions $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ are, in a sense, the most important, for we shall see in §3.4 that $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is, up to addition by a certain constant function, an automorphic integral of weight 2κ for Γ with a single pole at the cusp p in case Γ has width one at infinity (cf. §2.6) and $\kappa \leq 0$ and $m < 0$. We shall see also in §3.4 (cf. Theorems 3.4.3 and 3.4.4) that the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$, once corrected by adding a certain constant function in case $\kappa \leq 0$, is the expansion (of the correction by addition of a certain constant) of $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ at the cusp q in the sense of §2.7.

To further emphasize the importance of the functions $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ we observe that every Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is of the form $R^{\kappa(m)}_{\Gamma',\mathsf{p}'}(z)$ for some group Γ' with width one at infinity, and some cusp p' of Γ'. For if Σ_{p} and Σ_{q} are the chosen scaling sets for Γ at p and q, respectively, then we have $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = R^{\kappa(m)}_{U}(z)$ for $U = \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$. Recall (cf. (2.6.8)) that Γ^{q} is a shorthand for the group $\Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{q}}$, and this group Γ^{q} has width one at infinity by the defining properties of scaling cosets. If we define $\mathsf{p}^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1} \cdot \mathsf{p} \subset \hat{\mathbb{Q}}$ then p^{q} is a cusp of Γ^{q}, and $\Sigma_{\mathsf{p}^{\mathsf{q}}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}$ is a scaling coset for Γ^{q} at p^{q}. We now have $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} = \Sigma_{\mathsf{p}^{\mathsf{q}}}^{-1}$, so that $U = \llbracket\Sigma_{\mathsf{p}^{\mathsf{q}}}^{-1}$. Since Γ^{q} has width one at infinity we may take it to be a scaling coset for itself at the infinite cusp, and we thus have

$$R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = R^{\kappa(m)}_{\Gamma^{\mathsf{q}},\mathsf{p}^{\mathsf{q}}}(z) \tag{3.1.12}$$

subject to a consistent choice of scaling cosets for Γ and Γ^{q}. More explicitly, the equality in (3.1.12) holds in case $\Gamma^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{q}}$ and $\mathsf{p}^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1} \cdot \mathsf{p}$ and $\Sigma_{\mathsf{p}^{\mathsf{q}}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}$.

Our primary interest in this article is in the distinguished case that $\kappa = 0$. In order to simplify notation, and maintain consistency with the notation of §1, we write $R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ as a shorthand for $R^{0(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$.

We conclude this section with a result which expresses the Rademacher component function $R^{\kappa(m)}_{\llbracket\chi}(z)$ in terms of the functions $P^{\kappa(m)}_{\llbracket\chi}(z)$ of (3.1.1) and the integral operator J^{κ}_{w} of §2.7.

Lemma 3.1.1. *Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$, and let $\llbracket\chi \in \llbracket G(\mathbb{Q})$. Then we have*

$$R^{\kappa(m)}_{\llbracket\chi}(z) = P^{\kappa(m)}_{\llbracket\chi}(z) - m^{1-2\kappa}\left(J^{\kappa}_{\infty\cdot\llbracket\chi}P^{1-\kappa(-m)}_{\llbracket\chi}\right)(z). \tag{3.1.13}$$

Proof. In case $\llbracket\chi \in \llbracket G(\mathbb{Q})_{\infty}$ the second term in the right hand side of (3.1.13) vanishes (cf. (3.1.15)), and the identity (3.1.13) then holds by the definition of $R^{\kappa(m)}_{\llbracket\chi}(z)$, so we assume henceforth that $\llbracket\chi \in \llbracket G(\mathbb{Q})^{\times}$. By (3.1.8) we may write the Rademacher component function $R^{\kappa(m)}_{\llbracket\chi}(z)$ as

$$\mathrm{e}(m\llbracket\chi\cdot z)\,\mathrm{jac}(\llbracket\chi, z)^{\kappa} - \mathrm{e}(m\llbracket\chi\cdot\infty)\,\mathrm{e}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty)_{<1-2\kappa}\,\mathrm{jac}(\llbracket\chi, z)^{\kappa}. \tag{3.1.14}$$

The first term in (3.1.14) is just $P^{\kappa(m)}_{\llbracket\chi}(z)$, and a contour integral calculation confirms that the second term in (3.1.14) is $m^{1-2\kappa}$ times the image of $P^{1-\kappa(-m)}_{\llbracket\chi}(z)$ under $J^{\kappa}_{\infty\cdot\llbracket\chi}$. The details of this contour integral calculation are as follows. We have

$$\overline{\left(J^{\kappa}_{\infty\cdot\llbracket\chi}P^{1-\kappa(-m)}_{\llbracket\chi}\right)}(z) = (2\pi\mathbf{i})^{1-2\kappa}\int_{\chi^{-1}\rrbracket\cdot\infty}^{\infty}\mathrm{e}(-m\llbracket\chi\cdot\xi)\,\mathrm{jac}(\llbracket\chi, \xi)^{1-\kappa}(\xi - \bar{z})^{(-2\kappa)}\mathrm{d}\xi \tag{3.1.15}$$

by the definition of J_w^κ (cf. §2.7). Choosing a representative χ for the coset $\llbracket\chi$ we compute

$$\begin{aligned}
&\int_{\chi^{-1}\rrbracket\cdot\infty}^{\infty} \mathrm{e}(-m\llbracket\chi\cdot\xi)\,\mathrm{jac}(\llbracket\chi\,,\xi)^{1-\kappa}(\xi-\bar{z})^{(-2\kappa)}\mathrm{d}\xi \\
&= \mathrm{e}(-m\chi\cdot\infty)\int_{\infty}^{\chi\cdot\infty} \mathrm{e}(m\chi\cdot\infty-m\chi\cdot\xi)(\chi\cdot\xi-\chi\cdot\bar{z})^{(-2\kappa)}\mathrm{d}(\chi\cdot\xi)\,\mathrm{jac}(\chi,\bar{z})^{\kappa} \\
&= -\,\mathrm{e}(-m\chi\cdot\infty)\int_{\chi\cdot\infty}^{\infty} \mathrm{e}(m\chi\cdot\infty-m\xi)(\xi-\chi\cdot\bar{z})^{(-2\kappa)}\mathrm{d}\xi\,\mathrm{jac}(\chi,\bar{z})^{\kappa} \\
&= -\,\mathrm{e}(-m\chi\cdot\infty)\int_{0}^{\infty} \mathrm{e}(-m\xi)(\xi+\chi\cdot\infty-\chi\cdot\bar{z})^{(-2\kappa)}\mathrm{d}\xi\,\mathrm{jac}(\chi,\bar{z})^{\kappa}
\end{aligned} \tag{3.1.16}$$

where the last integral in (3.1.16) is taken over the vertical line $\{\mathbf{i}t \mid t\in\mathbb{R}^+\}$. Thus, making the substitution $2\pi\mathbf{i}m\xi=t$, we find that

$$\begin{aligned}
&\int_{\chi^{-1}\rrbracket\cdot\infty}^{\infty} \mathrm{e}(-m\llbracket\chi\cdot\xi)\,\mathrm{jac}(\llbracket\chi\,,\xi)^{1-\kappa}(\xi-\bar{z})^{(-2\kappa)}\mathrm{d}\xi \\
&= -\,\mathrm{e}(-m\chi\cdot\infty)(2\pi\mathbf{i}m)^{-1}\int_{0}^{\infty} e^{-t}((2\pi\mathbf{i}m)^{-1}t+\chi\cdot\infty-\chi\cdot\bar{z})^{(-2\kappa)}\mathrm{d}t\,\mathrm{jac}(\chi,\bar{z})^{\kappa} \\
&= -\,\mathrm{e}(-m\chi\cdot\infty)\sum_{k=0}^{-2\kappa}(2\pi\mathbf{i}m)^{2\kappa-1}(2\pi\mathbf{i}(m\chi\cdot\infty-m\chi\cdot\bar{z}))^{(k)}\,\mathrm{jac}(\chi,\bar{z})^{\kappa},
\end{aligned} \tag{3.1.17}$$

and from this we deduce that

$$\left(J^{\kappa}_{\infty\cdot\llbracket\chi}P^{1-\kappa(-m)}_{\llbracket\chi}\right)(z) = m^{2\kappa-1}\,\mathrm{e}(m\chi\cdot\infty)\,\mathrm{e}(m\chi\cdot z-m\chi\cdot\infty)_{<1-2\kappa}\,\mathrm{jac}(\chi,z)^{\kappa}. \tag{3.1.18}$$

Upon comparison with (3.1.14) we obtain the required identity (3.1.13). This completes the proof. □

3.2 Coefficients

For $\llbracket\chi\rrbracket\in\llbracket G(\mathbb{Q})\rrbracket^{\times}$ and $m,n\in\mathbb{Z}$ we define a holomorphic function $s\mapsto \mathrm{Kl}^s_{\llbracket\chi\rrbracket}(m,n)$ on $\mathbb{C}$ by setting

$$\mathrm{Kl}^s_{\llbracket\chi\rrbracket}(m,n) = \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\,\mathrm{e}(-n\left\llbracket\chi^{-1}\right\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket^{s}\,. \tag{3.2.1}$$

Here $\varrho\llbracket\chi\rrbracket$ denotes the scaling factor of (2.3.11) (cf. Lemma 2.4.1). If $\chi\in G(\mathbb{Q})^{\times}$ has a preferred representative $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ with $c>0$, then we have

$$\mathrm{Kl}^s_{\llbracket\chi\rrbracket}(m,n) = \mathrm{e}\left(\frac{ma+nd}{c}\right)\frac{(ad-bc)^s}{c^{2s}}, \tag{3.2.2}$$

so that sums of the $\mathrm{Kl}^s_{[\![\chi]\!]}(m,n)$ over suitable subsets of $[\![G(\mathbb{Q})]\!]^\times$ recover zeta functions of various kinds, as we shall see presently. With this in mind consider the series $\mathrm{Kl}^s_S(m,n)$ defined for S a subset of $[\![G(\mathbb{Q})]\!]^\times$ by setting

$$\mathrm{Kl}^s_S(m,n) = \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \mathrm{Kl}^s_{[\![\chi]\!]}(m,n) \tag{3.2.3}$$

where $S_{\leq K}$ is as in (2.4.7). Write $\mathrm{Kl}^s_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)$ for $\mathrm{Kl}^s_S(m,n)$ in the case that $S = \left[\!\left[\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\right]\!\right]^\times$ for Γ a subgroup of $G(\mathbb{Q})$ that is commensurable with $G(\mathbb{Z})$ and $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ is a system of scaling cosets for Γ. We will verify presently (cf. Proposition 3.2.1) that the Dirichlet series $s \mapsto \mathrm{Kl}^s_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)$ converges absolutely for $\Re(s) > 1$ (so that the limit is unnecessary there) and therefore defines a holomorphic function in this region; these functions $\mathrm{Kl}^s_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)$ are among the *Selberg–Kloosterman zeta functions* introduced by Selberg in [Sel65] and shown there to admit meromorphic continuation to the entire complex plane. Selberg demonstrated further in loc. cit. that the eigenfunctions of the hyperbolic Laplacian $\Delta = (z-\bar{z})^2\partial_z\partial_{\bar{z}}$ on $\Gamma\backslash\mathbb{H}$ determine the poles of $\mathrm{Kl}^s_S(m,n)$ in such a way that a pole at s corresponds to an eigenfunction with eigenvalue $\lambda = s(1-s)$. The convergence of expressions very similar to (3.2.3) in certain regions $\Re(s) \geq \sigma$ with $\sigma \leq 1$ was considered by Knopp in [Kno86] and [Kno89][1].

Proposition 3.2.1. *Let $\Gamma < G(\mathbb{Q})$ be a group commensurable with $G(\mathbb{Z})$, let $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ be a system of scaling cosets for Γ and set $S = \left[\!\left[\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\right]\!\right]^\times$ for some $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$. Then the series (3.2.3) defining $\mathrm{Kl}^s_S(m,n)$ converges for $s = 1$ when m and n are not both zero, and converges absolutely and locally uniformly in $\Re(s) > 1$ for any $m,n \in \mathbb{Z}$.*

Proof. For positive integers r and c let S^r_c denote the subset of S consisting of double cosets $[\![\chi]\!] \in [\![G(\mathbb{Q})]\!]^\times$ with $P\det(\chi) = r$ and $c(\chi) = c$, so that $\varrho(\chi) = r/c^2$ (cf. (2.3.6)) for $[\![\chi]\!] \in S^r_c$. Observe that there are at most c elements in S^r_c so we have

$$|\,\mathrm{Kl}^s_{S^r_c}(m,n)| \leq \sum_{[\![\chi]\!]\in S^r_c} |\,\mathrm{Kl}^s_{[\![\chi]\!]}(m,n)| \leq r^s c^{1-2s} \tag{3.2.4}$$

for $s \in \mathbb{C}$. Now $S = \bigcup_{r,c>0} S^r_c$ and there are only finitely many r for which $S^r = \bigcup_{c>0} S^r_c$ is non-empty since Γ is commensurable with $G(\mathbb{Z})$, so (3.2.4) and the identity $\mathrm{Kl}^s_S(m,n) = \sum_{c>0}\sum_{r>0} \mathrm{Kl}^s_{S^r_c}(m,n)$ implies that the series of absolute values obtained from (3.2.3) is bounded above by a constant independent of m and n so long as $\Re(s) > 1$. Indeed, we have

$$|\,\mathrm{Kl}^s_S(m,n)| \leq \sum_{c>0} \sum_{r\in P\det(S)} r^\sigma c^{1-2\sigma} \leq \#P\det(S) r^\sigma_{\max} \frac{1}{2\sigma-2} \tag{3.2.5}$$

---

[1] Please note that the latter of these references contains important corrections to the former.

where $\#P\mathrm{det}(S)$ is the number of values of r for which S^r is non-empty, $r_{\max}$ is the maximum r for which S^r is non-empty, and $\sigma = \Re(s)$. This proves the required convergence of (3.2.3) in the region $\Re(s) > 1$.

We now consider convergence at $s = 1$. As mentioned above, Selberg demonstrated the meromorphic continuation of the Selberg–Kloosterman zeta function $s \to \mathrm{Kl}_S^s(m,n)$ and related its poles to the spectrum of the Laplacian on $\Gamma\backslash\mathbb{H}$ in [Sel65]. It turns out that those poles in $\Re(s) > 1/2$ are finite in number and are confined to the real line segment $(1/2, 1)$ in the case that m and n are not both zero, and the absence of a pole at $s = 1$ derives from the fact that the eigenfunctions for Δ with eigenvalue 0 are harmonic, and thus constant. (See [Iwa02], and Chapter 9 especially, for a thorough development of these ideas.) In particular, if we write $Z_{m,n}(s)$ for Selberg's meromorphic continuation of $\mathrm{Kl}_S^s(m,n)$, then $Z_{m,n}(s)$ is holomorphic at s when $\Re(s) = 1$ and m and n are not both zero, and the only pole of $Z_{0,0}(s)$ on the line $\Re(s) = 1$ is at $s = 1$. On the other hand Goldfeld–Sarnak determined the estimate

$$|Z_{m,n}(s)| = O\left(|m||n|\frac{|s|^{1/2}}{\Re(s) - 1/2}\right) \tag{3.2.6}$$

for $mn \neq 0$ as $\Im(s) \to \infty$ with $\Re(s) > 1/2$ in [GS83]. (Loc. cit. considers the case that $m, n > 0$. See [Pri00] for the extension to $mn \neq 0$, and an improvement upon (3.2.6).) Armed with this bound and the holomorphy of $Z_{m,n}(s)$ in $\Re(s) > 1 - a$ for some $a > 0$ we can deduce the convergence of (3.2.3) at $s = 1$ in the following way, which is an adaptation of the argument presented in §2 of [Kno86]. Note that we can choose $a > 0$ such that $Z_{m,n}(s)$ has no poles in $(1-a, 1)$ for all $m, n \in \mathbb{Z}$ since the values of s such that $\Re(s) > 1/2$ and $s(1-s)$ is an eigenvalue of Δ are finite in number. For such a the estimate (3.2.6) holds uniformly in s for $\Re(s) > 1 - a$.

Let $a > 0$ such that the interval $(1-a, 1)$ contains no poles of $Z_{m,n}(s)$ for any $m, n \in \mathbb{Z}$, let $T > 0$ and set C to be the positively oriented rectangular contour with corners $\pm a \pm \mathbf{i}T$. Consider the contour integral

$$I = \frac{1}{2\pi\mathbf{i}} \int_C Z_{m,n}\left(\frac{t}{2} + 1\right) x^t \frac{\mathrm{d}t}{t} \tag{3.2.7}$$

where $x > 0$. By our choice of a the integrand has a unique simple pole inside C and the residue there is $Z_{m,n}(1)$, so $I = Z_{m,n}(1)$. On the other hand the integral over the right most vertical segment of C contributes $\Sigma_T(x)$ where

$$\Sigma_T(x) = \frac{1}{2\pi\mathbf{i}} \int_{a-\mathbf{i}T}^{a+\mathbf{i}T} Z_{m,n}\left(\frac{t}{2} + 1\right) x^t \frac{\mathrm{d}t}{t}. \tag{3.2.8}$$

Observe that in the limit as $T \to \infty$ we obtain the function

$$\Sigma(x) = \lim_{T\to\infty} \Sigma_T(x) = \sum_{r>0} \sum_{0<c<x\sqrt{r}} \mathrm{Kl}^1_{S^r_c}(m,n) \tag{3.2.9}$$

(so long as x does not coincide with $c/\sqrt{r}$ for some $c, r > 0$ with S^r_c non-empty) and the difference $|\Sigma(x) - \Sigma_T(x)|$ is $O(x^a T^{-1})$ where the implied constant is independent of m and n. The estimate (3.2.6) implies that the contribution to I of the left most vertical segment of C is $O(|m||n|x^{-a}T^{1/2})$, and we may show that the contributions of the two horizontal components of C are $O(|m||n|x^a T^{1/4-1/4a})$ so long as $T < x^{4a}$ by applying the Phragmén–Lindelöf theorem (as stated in Theorem 14 of [HR64]) to the estimates $|Z_{m,n}(t/2+1)| = O(|m||n||\Im(t)|^{1/2})$ for $\Re(t) = -a$ and $|Z_{m,n}(t/2+1)| = O(1)$ for $\Re(t) = a$ (where the former of these is another application of (3.2.6) and the constant implied by the latter is independent of both m and n according to Proposition 3.2.1). Setting $T = x^{3a/2}$ we obtain

$$|\Sigma(x) - Z_{m,n}(1)| = O(|m||n|x^{-a/4}) \tag{3.2.10}$$

under the condition that $a < 3/13$, and in the limit as $x \to \infty$ we arrive at the convergence of the sum (3.2.3) at $s = 1$ in the case that $mn \neq 0$.

For the case that $mn = 0$ but $m + n \neq 0$ we note that a similar estimate to (3.2.6)

$$|Z_{m,n}(s)| = O\left(|s|^{1/2}\right) \tag{3.2.11}$$

holds uniformly in s for $\Re(s) > 1 - a$, with the implied constant independent of m and n, as a consequence of the boundedness of Fourier coefficients of non-analytic Eisenstein series in this region (cf. [Kub73]). Now the argument of the preceding paragraph goes through with (3.2.11) used in place of (3.2.6) and we obtain the desired convergence of $\mathrm{Kl}^s_S(m,n)$ at $s = 1$ for all $m, n \in \mathbb{Z}$ such that m and n are not both zero. This completes the proof of the proposition. □

Remark 3.2.2. Consider the case that $S = [\![G(\mathbb{Z})]\!]^\times$. Then the function $\mathrm{Kl}^s_S(m,n)$ is closely related to the Riemann zeta function $\zeta(s)$, for we have

$$\mathrm{Kl}^s_S(m,0) = \sum_{c\in\mathbb{Z}^+} \sum_{a\in(\mathbb{Z}/c)^*} \frac{\mathrm{e}(ma/c)}{c^{2s}} = \frac{\sigma(|m|, 1-2s)}{\zeta(2s)} \tag{3.2.12}$$

for $\Re(s) > 1$ in the case that $m < 0$, where $\sigma(n,s)$ is the divisor function of (2.2.1). In particular, the function $s \mapsto \mathrm{Kl}^{s/2}_S(1,0)$ is just the reciprocal of the Riemann zeta function when $S = [\![G(\mathbb{Z})]\!]^\times$. When $m = n = 0$ we have $\mathrm{Kl}^s_S(0,0) = \sum_{c>0} \phi(c)c^{-2s}$ where ϕ is Euler's totient function so clearly the sum diverges at $s = 1$ in this case. In general the residue of $\mathrm{Kl}^s_S(0,0)$ at $s = 1$ is the reciprocal of the hyperbolic area of the quotient $\Gamma\backslash\mathbb{H}$ when $S = [\![\Gamma]\!]^\times$ for some group Γ. (Cf. [Iwa02, §6.4].)

The following result is evident from the definition of $\mathrm{Kl}^{s}_{[\![\chi]\!]}(m,n)$, and the definition of the conjugation $\chi \mapsto \bar{\chi}$ (cf. §2.3).

Lemma 3.2.3. *For any $\chi \in G(\mathbb{Q})^{\times}$, $m,n \in \mathbb{Z}$, and $s \in \mathbb{C}$, we have*

$$\mathrm{Kl}^{s}_{[\![\chi]\!]}(m,n) = \mathrm{Kl}^{s}_{[\![\chi^{-1}]\!]}(n,m) = \mathrm{Kl}^{s}_{[\![\bar{\chi}]\!]}(-m,-n), \tag{3.2.13}$$

$$\mathrm{Kl}^{s}_{[\![\chi]\!]}(m,n)\varrho\,[\![\chi]\!]^{1-s} = \mathrm{Kl}^{1-s}_{[\![\chi]\!]}(m,n)\varrho\,[\![\chi]\!]^{s}. \tag{3.2.14}$$

For a continuous analogue of the function $\mathrm{Kl}^{s}_{[\![\chi]\!]}(m,n)$ we define $\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z)$, for $\chi \in G(\mathbb{Q})^{\times}$, $w,z \in \mathbb{C}$ and $\kappa \in \mathbb{Z}$, by setting

$$\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z) = 2\pi\mathbf{i}\,\mathrm{Res}_{\xi=0}\,\mathrm{e}\left(\frac{w\varrho\,[\![\chi]\!]}{\xi}\right)\mathrm{e}(z\xi)\,\frac{1}{\xi^{2\kappa}}\,\mathrm{d}\xi. \tag{3.2.15}$$

Lemma 3.2.4. *The function $\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z)$ admits the following series representations.*

$$\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z) = \sum_{k\geq 0}(2\pi\mathbf{i})^{2k+2\kappa}\varrho\,[\![\chi]\!]^{k}\,w^{(k)}z^{(k+2\kappa-1)} \tag{3.2.16}$$

$$\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z) = \sum_{k\geq 0}(2\pi\mathbf{i})^{2k+2-2\kappa}\varrho\,[\![\chi]\!]^{k+1-2\kappa}\,w^{(k+1-2\kappa)}z^{(k)} \tag{3.2.17}$$

The first expression (3.2.16) holds for $\kappa > 0$. The second expression (3.2.17) holds for $\kappa \leq 0$.

Lemma 3.2.5. *For any $\chi \in G(\mathbb{Q})^{\times}$ and $w,z \in \mathbb{C}$ and $\kappa \in \mathbb{Z}$ we have*

$$\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(w,z)\varrho\,[\![\chi]\!]^{\kappa} = -\,\mathrm{Bl}^{1-\kappa}_{[\![\chi]\!]}(z,w)\varrho\,[\![\chi]\!]^{1-\kappa} = -\,\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(-w,-z)\varrho\,[\![\chi]\!]^{\kappa}. \tag{3.2.18}$$

The main construction of this section is the following. Given a subset $S \subset [\![G(\mathbb{Q})]\!]^{\times}$ and $\kappa \in \mathbb{Z}$, we define the *coefficient function*, denoted $(m,n) \mapsto \mathrm{c}^{\kappa}_{S}(m,n)$, by setting

$$\mathrm{c}^{\kappa}_{S}(m,n) = \lim_{K\to\infty}\sum_{[\![\chi]\!]\in S_{\leq K}}\mathrm{Kl}^{\kappa}_{[\![\chi]\!]}(m,n)\,\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n) \tag{3.2.19}$$

(cf. (2.4.7)) whenever this limit converges. In applications we take S to be of the form $S = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ where Σ_{p} and Σ_{q} are scaling cosets for a group Γ at cusps $\mathsf{p},\mathsf{q} \in \mathsf{P}_{\Gamma}$ where Γ is commensurable with $G(\mathbb{Z})$. We will see presently that $\mathrm{c}^{\kappa}_{S}(m,n)$ converges for any such S. We obtain the following formulas by combining (3.2.1) and Lemma 3.2.4.

Lemma 3.2.6. *The coefficient functions* $\mathrm{c}^{\kappa}_{S}(m,n)$ *admit the following series representations.*

$$\mathrm{c}^{\kappa}_{S}(m,n) = \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \sum_{k\geq 0} \mathrm{e}(m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}(-n\,[\![\chi^{-1}]\!]\cdot\infty)(-4\pi^2\varrho\,[\![\chi]\!])^{k+\kappa} m^{(k)} n^{(k+2\kappa-1)} \tag{3.2.20}$$

$$\mathrm{c}^{\kappa}_{S}(m,n) = \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \sum_{k\geq 0} \mathrm{e}(m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}(-n\,[\![\chi^{-1}]\!]\cdot\infty)(-4\pi^2\varrho\,[\![\chi]\!])^{k+1-\kappa} m^{(k+1-2\kappa)} n^{(k)} \tag{3.2.21}$$

The expression (3.2.20) holds for $\kappa > 0$. *The expression (3.2.21) holds for* $\kappa \leq 0$.

Proposition 3.2.7. *For* $S \subset [\![G(\mathbb{Q})]\!]^{\times}$ *and* $m,n,\kappa\in\mathbb{Z}$ *we have*

$$\mathrm{c}^{\kappa}_{S}(m,n) = -\mathrm{c}^{1-\kappa}_{S^{-1}}(n,m) = -\mathrm{c}^{\kappa}_{\bar{S}}(-m,-n). \tag{3.2.22}$$

Proof. These identities follow directly from Lemma 3.2.6, and the observation that the values $\varrho\,[\![\chi]\!]$, $\varrho\,[\![\chi^{-1}]\!]$ and $\varrho\,[\![\bar{\chi}]\!]$ all coincide for any $\chi\in G(\mathbb{Q})^{\times}$. ☐

Proposition 3.2.8. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$ *and let* $\mathsf{p},\mathsf{q}\in\mathsf{P}_{\Gamma}$ *be cusps of* Γ. *Then the expression (3.2.19) defining the coefficient function* $\mathrm{c}^{\kappa}_{S}(m,n)$ *converges for* $S = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ *and any* $\kappa,m,n\in\mathbb{Z}$, *and further, we have*

$$|\mathrm{c}^{\kappa}_{S}(m,n)| = O\left(|m|^{1/2-\kappa}|n|^{\kappa-1/2}e^{C|mn|^{1/2}}\right) \tag{3.2.23}$$

as $|n|\to\infty$ *where* C *depends only on* S *and the implied constant depends only on* S *and* κ.

Proof. Observe that according to the $\kappa\mapsto 1-\kappa$ symmetry of Proposition 3.2.7 it suffices to consider the case that $\kappa > 0$. Observe also that for $\kappa > 0$ the coefficient function $\mathrm{c}^{\kappa}_{S}(m,n)$ vanishes for all m when $n = 0$, and when $m = 0$ we have

$$\mathrm{c}^{\kappa}_{S}(0,n) = (-4\pi^2)^{\kappa} n^{(2\kappa-1)}\,\mathrm{Kl}^{\kappa}_{S}(0,n) \tag{3.2.24}$$

according to (3.2.20), so the convergence in case $mn = 0$ is either trivial or follows from the convergence of the Selberg–Kloosterman zeta function $\mathrm{Kl}^{\kappa}_{S}(m,n)$ which is established in Proposition 3.2.1. Assume then that $mn\neq 0$. By (3.2.20) we have

$$\mathrm{c}^{\kappa}_{S}(m,n) = (-4\pi^2)^{\kappa}\sum_{c>0}\sum_{r>0}\sum_{k\geq 0}\mathrm{Kl}^{\kappa+k}_{S^{r}_{c}}(m,n)(-4\pi^2 m)^{(k)} n^{(k+2\kappa-1)} \tag{3.2.25}$$

where S^{r}_{c} is defined as in the proof of Proposition 3.2.1 and the sum over r is finite. In case

$\kappa > 1$ we can reorder the summation over k in (3.2.25) past the other two since the sum

$$\mathrm{c}^{\kappa}_{S}(m,n) = (-4\pi^2)^{\kappa} \sum_{k\geq 0} (-4\pi^2 m)^{(k)} n^{(k+2\kappa-1)} \operatorname{Kl}^{\kappa+k}_{S}(m,n) \tag{3.2.26}$$

so obtained is absolutely convergent according to the estimate (3.2.5). In detail,

$$\begin{aligned} |\mathrm{c}^{\kappa}_{S}(m,n)| &\leq (4\pi^2)^{\kappa} \sum_{k\geq 0} (4\pi^2|m|)^{(k)} |n|^{(k+2\kappa-1)} \#P\det(S) r_{\max}^{\kappa+k} \frac{1}{2\kappa-2+2k} \\ &\leq \#P\det(S)(4\pi^2 r_{\max})^{\kappa} \sum_{k\geq 0} (4\pi^2 r_{\max}|m|)^{(k)} |n|^{(k+2\kappa-1)} < \infty. \end{aligned} \tag{3.2.27}$$

Now for A and B non-negative we have

$$\sum_{k\geq 0} A^{(k)} B^{(k+2\kappa-1)} \leq A^{1/2-\kappa} B^{\kappa-1/2} \sum_{k\geq 0} (2(AB)^{1/2})^{(2k+2\kappa-1)} \tag{3.2.28}$$

so that setting $A = 4\pi^2 r_{\max}|m|$ and $B = |n|$ in (3.2.28) we obtain the required estimate (3.2.23) from (3.2.27) with $C = 4\pi r_{\max}^{1/2}$.

It remains to establish the estimate (3.2.23) in case $\kappa = 1$. For this we write

$$\mathrm{c}^{\kappa}_{S}(m,n) = (-4\pi^2)^{\kappa} n^{(2\kappa-1)} \operatorname{Kl}^{\kappa}_{S}(m,n) + \mathrm{c}^{\kappa}_{S}(m,n)_{+} \tag{3.2.29}$$

where

$$\mathrm{c}^{\kappa}_{S}(m,n)_{+} = (-4\pi^2)^{\kappa} \sum_{k>0} (-4\pi^2 m)^{(k)} n^{(k+2\kappa-1)} \operatorname{Kl}^{\kappa+k}_{S}(m,n) \tag{3.2.30}$$

is an absolutely convergent sum (cf. (3.2.26)) and we have

$$|\mathrm{c}^{\kappa}_{S}(m,n)_{+}| = O\left(|m|^{1/2-\kappa}|n|^{\kappa-1/2} e^{C|mn|^{1/2}}\right) \tag{3.2.31}$$

with $C = 4\pi r_{\max}^{1/2}$ according to the estimates (3.2.27) and (3.2.28). Now we see from (3.2.29) that the estimate (3.2.31) also holds for $\mathrm{c}^{\kappa}_{S}(m,n)$ since (3.2.6) (or (3.2.11) in case $m = 0$) implies that $\operatorname{Kl}^{\kappa}_{S}(m,n)$ has at most polynomial growth in n as $|n| \to \infty$ when $\kappa = 1$. This completes the proof. □

We write $\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)$ for $\mathrm{c}^{\kappa}_{S}(m,n)$ in the case that $S = \left[\!\left[\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\right]\!\right]^{\times}$. This notation suppresses the dependence on the choice of scaling cosets; the following lemma describes this dependence explicitly.

Lemma 3.2.9. *Let* $\mathsf{p},\mathsf{q} \in \mathsf{P}_{\Gamma}$*, let* Σ_{p} *and* Σ'_{p} *be scaling cosets for* Γ *at* p *and let* Σ_{q} *and* Σ'_{q} *be scaling cosets for* Γ *at* q*. Set* $S = \left[\!\left[\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\right]\!\right]^{\times}$ *and* $S' = \left[\!\left[(\Sigma'_{\mathsf{p}})^{-1}\Sigma'_{\mathsf{q}}\right]\!\right]^{\times}$*. Then we have* $\Sigma'_{\mathsf{p}} = \Sigma_{\mathsf{p}} T^{\alpha}$

and $\Sigma'_{\mathsf{q}} = \Sigma_{\mathsf{q}} T^{\beta}$ *and* $\mathrm{c}^{\kappa}_{S'}(m,n) = \mathrm{e}(m\alpha + n\beta)\mathrm{c}^{\kappa}_{S}(m,n)$ *for some* $\alpha, \beta \in \mathbb{Q}$.

With a fixed choice of subset $S \subset [\![G(\mathbb{Q})]\!]^{\times}$ and integer $\kappa \in \mathbb{Z}$, we assemble the coefficient functions $\mathrm{c}^{\kappa}_{S}(m,n)$ into a formal series $\tilde{F}^{\kappa}_{S}(p,q)$ by setting

$$\tilde{F}^{\kappa}_{S}(p,q) = \sum_{m,n\in\mathbb{Z}} \mathrm{c}^{\kappa}_{S}(m,n)p^m q^n. \tag{3.2.32}$$

The formal Fourier coefficients of the formal series $\tilde{F}^{\kappa}_{S}(p,q)$ with respect to the variable p are of particular importance. For fixed $S \subset [\![G(\mathbb{Q})]\!]^{\times}$ and $\kappa \in \mathbb{Z}$ we define formal series $\tilde{F}^{\kappa(m)}_{S}(q)$ by requiring that

$$\tilde{F}^{\kappa}_{S}(p,q) = \sum_{m\in\mathbb{Z}} \tilde{F}^{\kappa(m)}_{S}(q)p^m. \tag{3.2.33}$$

We define the *singular* and *regular part of* $\tilde{F}^{\kappa(m)}_{S}(q)$, to be denoted $\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{sing}}$ and $\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{reg}}$, respectively, by setting

$$\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{sing}} = \sum_{n\geq 0} \mathrm{c}^{\kappa}_{S}(m,-n-1)q^{-n-1}, \quad \tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{reg}} = \sum_{n\geq 0} \mathrm{c}^{\kappa}_{S}(m,n)q^{n}. \tag{3.2.34}$$

We define $\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{van}}$ to be the series obtained by removing the constant term from $\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{reg}}$.

$$\tilde{F}^{\kappa(m)}_{S}(q)_{\mathrm{van}} = \sum_{n>0} \mathrm{c}^{\kappa}_{S}(m,n)q^{n} \tag{3.2.35}$$

Given the formal series $\tilde{F}^{\kappa}_{S}(p,q)$ for some S and κ, one may ask if the coefficient series $\tilde{F}^{\kappa(m)}_{S}(q)$ define holomorphic functions on $\mathbb{H}$ upon the substitution of $\mathrm{e}(z)$ for q. Supposing this to be the case, we define $F^{\kappa(m)}_{S}(z)$ by setting

$$F^{\kappa(m)}_{S}(z) = \tilde{F}^{\kappa(m)}_{S}(\mathrm{e}(z)) = \sum_{n\in\mathbb{Z}} \mathrm{c}^{\kappa}_{S}(m,n)\,\mathrm{e}(nz), \tag{3.2.36}$$

and we define $F^{\kappa(m)}_{S}(z)_{\mathrm{sing}}$, $F^{\kappa(m)}_{S}(z)_{\mathrm{reg}}$ and $F^{\kappa(m)}_{S}(z)_{\mathrm{van}}$ in the analogous way.

We write $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for $F^{\kappa(m)}_{S}(z)$ in the case that $S = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ for Γ a group commensurable with $G(\mathbb{Z})$ and for $\Sigma_{\mathsf{p}}, \Sigma_{\mathsf{q}} \in \Gamma\backslash G(\mathbb{Q})$ scaling cosets for Γ at cusps $\mathsf{p}, \mathsf{q} \in \mathsf{P}_{\Gamma}$. We write $F^{(m)}_{S}(z)$ as a shorthand for $F^{0(m)}_{S}(z)$, and interpret the notation $F^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$, &c., similarly.

The regular part of the formal series $\tilde{F}^{\kappa(m)}_{S}(q)$ indeed defines a holomorphic function on $\mathbb{H}$, for suitable $\kappa, m \in \mathbb{Z}$, in the case that S is of the form $[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$.

Proposition 3.2.10. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$*, let* $\mathsf{p}, \mathsf{q} \in \mathsf{P}_{\Gamma}$ *be cusps of* Γ

and let $\kappa, m \in \mathbb{Z}$. Then the series

$$F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}} = \sum_{n>0} \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)\,\mathrm{e}(nz) \tag{3.2.37}$$

converges absolutely and locally uniformly for $z \in \mathbb{H}$. In particular, $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}$ is a holomorphic function on $\mathbb{H}$ for any $\kappa, m \in \mathbb{Z}$.

Proof. For $m \neq 0$ the claim follows directly from the estimate (3.2.23). If $m = 0$ and $\kappa \leq 0$ then $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}$ vanishes identically according to (3.2.21). For $m = 0$ and $\kappa > 0$ the required convergence of (3.2.37) follows from (3.2.24) since $\mathrm{Kl}^{\kappa}_{S}(0,n)$ has at most polynomial growth in n according to the proof of Proposition 3.2.1. □

By a similar argument to that just given we see that the singular part of the formal series $\tilde{F}^{\kappa(m)}_{S}(q)$ defines an anti-holomorphic function on $\mathbb{H}$ when S is of the form $[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$.

Proposition 3.2.11. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_{\Gamma}$ be cusps of Γ and let $\kappa, m \in \mathbb{Z}$. Then the series*

$$F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(\bar{z})_{\mathrm{sing}} = \sum_{n>0} \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,-n)\,\mathrm{e}(-n\bar{z}) \tag{3.2.38}$$

converges absolutely and locally uniformly for $z \in \mathbb{H}$. In particular, $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(\bar{z})_{\mathrm{sing}}$ is an anti-holomorphic function on $\mathbb{H}$ for any $\kappa, m \in \mathbb{Z}$.

The notation $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}$ suppresses the dependence upon the choice of scaling cosets Σ_{p} and Σ_{q}. The next result encodes this dependence precisely.

Proposition 3.2.12. *Let Γ be a group commensurable with $G(\mathbb{Z})$, and let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_{\Gamma}$. Let Σ_{p} and Σ'_{p} be scaling cosets for Γ at p and let Σ_{q} and Σ'_{q} be scaling cosets for Γ at q. Set $S = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ and $S' = [\![(\Sigma'_{\mathsf{p}})^{-1}\Sigma'_{\mathsf{q}}]\!]^{\times}$. Then we have $\Sigma'_{\mathsf{p}} = \Sigma_{\mathsf{p}}T^{\alpha}$ and $\Sigma'_{\mathsf{q}} = \Sigma_{\mathsf{q}}T^{\beta}$ and $F^{\kappa(m)}_{S'}(z) = \mathrm{e}(m\alpha)F^{\kappa(m)}_{S}(z+\beta)$ for some $\alpha, \beta \in \mathbb{Q}$.*

3.3 Convergence

Our main goal in this section is to establish the convergence, and Fourier series expansion, of the Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ (cf. §3.1), attached to a group Γ commensurable with $G(\mathbb{Z})$, and cusps p and q for Γ. We will require mostly minor modifications of the arguments furnished by Niebur in [Nie74]. An important technical tool in these arguments is the identity

$$\sum_{n>0} n^{(s-1)}\,\mathrm{e}(nz) = \sum_{n\in\mathbb{Z}} (-2\pi\mathbf{i})^{-s}(z+n)^{-s}, \tag{3.3.1}$$

valid for $\Re(s) > 1$, in which both sums converge absolutely and uniformly in z on compact subsets of $\mathbb{H}$. A nice proof of (3.3.1) appears in [KR01]. This identity can be extended to $s = 1$ at the expense of absolute convergence.

Lemma 3.3.1. *Let $z \in \mathbb{H}$ and let $K \in \mathbb{Z}^+$ such that $K + 1/2 > \Re(z)$. Then we have*

$$\sum_{n>0} \mathrm{e}(nz) = \sum_{\substack{n\in\mathbb{Z} \\ -K\leq n\leq K}} (-2\pi\mathbf{i})^{-1}(z+n)^{-1} + \lambda_K(z) \tag{3.3.2}$$

where the error term $\lambda_K(z)$ is given by

$$\lambda_K(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{z+\mathbf{i}v-K-1/2} - \frac{1}{z+\mathbf{i}v+K+1/2}\right)\frac{1}{1+e^{-2\pi v}}\mathrm{d}v. \tag{3.3.3}$$

Proof. The proof follows that of (the more general) Lemma 4.1 in [Nie74]. Define a function $f(\xi)$ by setting $f(\xi) = (2\pi\mathbf{i}\xi)^{-1}(\mathrm{e}(\xi - z) - 1)^{-1}$. Then $f(\xi)$ has a pole at $\xi = 0$ and a pole also at $\xi = z + n$ for each $n \in \mathbb{Z}$. For the residue of $f(\xi)\mathrm{d}\xi$ at $\xi = 0$ we have

$$\mathrm{Res}_{\xi=0}\, f(\xi)\mathrm{d}\xi = \frac{1}{2\pi\mathbf{i}}\frac{1}{\mathrm{e}(-z)-1} = \frac{1}{2\pi\mathbf{i}}\sum_{n>0}\mathrm{e}(nz), \tag{3.3.4}$$

which is $1/2\pi\mathbf{i}$ times the left hand side of (3.3.2), while the residue at $\xi = z+n$ is $(2\pi\mathbf{i})^{-2}(z+n)^{-1}$. Let K be a positive integer such that $K + 1/2 > \Re(z)$ and set $Q = K + 1/2$ for convenience. Let γ be the positively oriented boundary of the rectangle with corners $\Re(z) \pm Q \pm \mathbf{i}L$ where $L > \Im(z)$. Then by the residue theorem we have

$$\int_\gamma f(\xi)\mathrm{d}\xi = \sum_{n>0}\mathrm{e}(nz) + \sum_{-K\leq n\leq K}(2\pi\mathbf{i})^{-1}(z+n)^{-1}. \tag{3.3.5}$$

On the other hand, the contributions of the horizontal edges of γ to $\int_\gamma f(\xi)\mathrm{d}\xi$ tend to 0 as $L \to \infty$ since $f(\xi)$ decays exponentially as $\Im(\xi) \to \infty$. The residues are independent of L for L sufficiently large so taking the limit as $L \to \infty$ we obtain

$$\sum_{n>0}\mathrm{e}(nz) + \sum_{-K\leq n\leq K}(2\pi\mathbf{i})^{-1}(z+n)^{-1} = \lambda_K(z) \tag{3.3.6}$$

where $\lambda_K(z)$ is the limit as $L \to \infty$ of the contributions of the vertical edges of γ. Setting $x = \Re(z)$ we have

$$\lambda_K(z) = \mathbf{i}\int_{-\infty}^{\infty}(f(x+Q+\mathbf{i}v) - f(x-Q+\mathbf{i}v))\mathrm{d}v. \tag{3.3.7}$$

After replacing v with $v + \Im(z)$ and observing that $\mathrm{e}(Q) = \mathrm{e}(-Q) = -1$ since K is an integer we obtain the expression

$$\lambda_K(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{z+\mathbf{i}v-Q}-\frac{1}{z+\mathbf{i}v+Q}\right)\frac{1}{1+\mathrm{e}(\mathbf{i}v)}\mathrm{d}v \tag{3.3.8}$$

and combining this with (3.3.6) recovers the required identity. □

Setting $Q = K + 1/2$ and substituting Qv for v in (3.3.3) we find that

$$\begin{aligned}\lambda_K(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{z/Q+\mathbf{i}v-1}-\frac{1}{z/Q+\mathbf{i}v+1}\right)\frac{1}{1+\mathrm{e}(\mathbf{i}Qv)}\mathrm{d}v\\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{((z/Q+\mathbf{i}v)^2-1)}\frac{1}{1+e^{-2\pi Qv}}\mathrm{d}v\end{aligned} \tag{3.3.9}$$

which demonstrates that $\lambda_K(z)$ tends to $-1/\pi$ times $\int_0^\infty(1+v^2)^{-1}\mathrm{d}v = \pi/2$ as $K\to\infty$. We thus obtain

$$\sum_{n>0}\mathrm{e}(nz) = -\frac{1}{2} + \lim_{K\to\infty}\sum_{-K<n<K}(-2\pi\mathbf{i})^{-1}(z+n)^{-1}. \tag{3.3.10}$$

The formulas (3.3.1) and (3.3.10) are collectively known as the *Lipschitz summation formula*.

Theorem 3.3.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$, and let $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m\in\mathbb{Z}$ such that $\kappa\le 0$ and $m<0$. Then the limit defining the classical Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ converges locally uniformly to a holomorphic function on $\mathbb{H}$, and we have*

$$R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + \frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0) + F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}. \tag{3.3.11}$$

Proof. Set $U = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ and $S = [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$. For $\kappa<0$ the result follows from Lemmas 4.2 and 4.3 of [Nie74]. Indeed, using Lemma 3.1.1 we may identify the Rademacher component function $R^{\kappa(m)}_{[\![\chi}(z)$ with the function $s_{[\![\chi}(z)+t_{[\![\chi}(z)$ defined on page 376 of [Nie74]. Then the expression

$$\sum_{c\in\mathbb{Z}^+}\lim_{K\to\infty}\sum_{\substack{[\![\chi\in U_{\le K}\\ c(\chi)=c}}R^{\kappa(m)}_{[\![\chi}(z) \tag{3.3.12}$$

is, up to addition by a certain constant function, the definition of Rademacher sum used by Niebur in loc. cit. The Lipschitz summation formula(s) (3.3.1) and (3.3.10) are the main tools used in equating (3.3.12) with the right hand side of (3.3.11).

In case $\kappa=0$ we require to reinforce the argument of [Nie74] with a non-trivial estimate for the Selberg–Kloosterman zeta function $\mathrm{Kl}^s_S(m,n)$ at $s=1$; we will apply the result (3.2.6) of

Goldfeld–Sarnark to this purpose. Assume then that $\kappa = 0$ and observe that we have

$$\begin{aligned} R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) &= \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + \lim_{K\to\infty} \sum_{\llbracket\chi\in U^{\times}_{\leq K}} \mathrm{e}(m\llbracket\chi\cdot\infty)\,\mathrm{e}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty, 1) \\ &= \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + \lim_{K\to\infty} \sum_{\llbracket\chi\in U^{\times}_{\leq K}} \mathrm{e}(m\llbracket\chi\cdot\infty)\sum_{k\geq 0}(-2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket)^{k+1}m^{(k+1)}(z-\chi^{-1}\rrbracket\cdot\infty)^{-k-1} \end{aligned} \tag{3.3.13}$$

according to (3.1.7), the definition (2.1.6) of the generalized exponential $\mathrm{e}(z,s)$, and the identity (2.4.4). We write $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = \mathrm{e}(mz) + R_0 + R_+$ where

$$R_0 = \lim_{K\to\infty} \sum_{\llbracket\chi\in U^{\times}_{\leq K}} \mathrm{e}(m\llbracket\chi\cdot\infty)(-2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket)m(z-\chi^{-1}\rrbracket\cdot\infty)^{-1}, \tag{3.3.14}$$

$$R_+ = \sum_{\llbracket\chi\in U^{\times}} \mathrm{e}(m\llbracket\chi\cdot\infty)\sum_{k>0}(-2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket)^{k+1}m^{(k+1)}(z-\chi^{-1}\rrbracket\cdot\infty)^{-k-1}, \tag{3.3.15}$$

and the second sum R_+ is absolutely convergent, locally uniformly for $z \in \mathbb{H}$. After choosing a representative χ for each double coset $\llbracket\chi\rrbracket$ in S we may rewrite R_+ as

$$R_+ = \sum_{\llbracket\chi\rrbracket\in S} \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\sum_{k>0}(-2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket)^{k+1}m^{(k+1)}\sum_{n\in\mathbb{Z}}(z-\chi^{-1}\rrbracket\cdot\infty+n)^{-k-1}, \tag{3.3.16}$$

and applying the Lipschitz summation formula (3.3.1) to this we obtain

$$\begin{aligned} R_+ &= \sum_{n>0}\sum_{\llbracket\chi\rrbracket\in S} \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\,\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)\sum_{k>0}(-4\pi^2\varrho\llbracket\chi\rrbracket)^{k+1}m^{(k+1)}n^{(k)}\,\mathrm{e}(nz) \\ &= \sum_{n>0}\mathrm{c}^{\kappa}_{S}(m,n)_+\,\mathrm{e}(nz) \end{aligned} \tag{3.3.17}$$

where $\mathrm{c}^{\kappa}_{S}(m,n)_+$ satisfies $\mathrm{c}^{\kappa}_{S}(m,n) = (-4\pi^2)m\,\mathrm{Kl}^{1}_{S}(m,n) + \mathrm{c}^{\kappa}_{S}(m,n)_+$ according to (3.2.21). For K a positive integer set

$$R_0(K) = (-2\pi\mathbf{i})m\sum_{U^{\times}_{\leq K}} \mathrm{e}(m\llbracket\chi\cdot\infty)\varrho\llbracket\chi\rrbracket\,(z-\chi^{-1}\rrbracket\cdot\infty)^{-1} \tag{3.3.18}$$

so that $R_0 = \lim_{K\to\infty} R_0(K)$. Then choosing as above a representative χ for each $\llbracket\chi\rrbracket \in S$ we have

$$\begin{aligned}R_0(K) &= (-2\pi\mathbf{i})m \sum_{\llbracket\chi\rrbracket\in S_{\leq K}} \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket \sum_{\substack{n\in\mathbb{Z}\\ -K^2\leq d(\chi)+nc(\chi)\leq K^2}} (z-\chi^{-1}\rrbracket\cdot\infty+n)^{-1}\\ &= (-2\pi\mathbf{i})m\sum_{0<c<K}\sum_{\llbracket\chi\rrbracket\in S_c}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket\left(\sum_{|n|\leq K^2/c}(z-\chi^{-1}\rrbracket\cdot\infty+n)^{-1}+O(c/K^2)\right)\end{aligned} \tag{3.3.19}$$

where the term $O(c/K^2)$ accounts for the discrepency between the two summations over n. Applying the Lipschitz summation formula (3.3.6) to the second line of (3.3.19) we obtain

$$R_0(K) = (-4\pi^2)m\sum_{0<c<K}\sum_{\llbracket\chi\rrbracket\in S_c}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket\left(\frac{1}{2}+\sum_{n>0}\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)\,\mathrm{e}(nz)+O(c/K^2)\right) \tag{3.3.20}$$

since the error term $\lambda_X(z)$ appearing in (3.3.6) satisfies $\lambda_X(z)+1/2=O(1/X)$. We conclude from this that

$$R_0 = \lim_{K\to\infty}\sum_{\llbracket\chi\rrbracket\in S_{\leq K}}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)(-4\pi^2)\varrho\llbracket\chi\rrbracket\, m\left(\frac{1}{2}+\sum_{n>0}\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)\,\mathrm{e}(nz)\right). \tag{3.3.21}$$

We claim now that $R_0 = \mathrm{c}^{\kappa}_S(m,0)/2+\sum_{n>0}\mathrm{c}^{\kappa}_S(m,n)_0\,\mathrm{e}(nz)$ where $\mathrm{c}^{\kappa}_S(m,n)_0 = \mathrm{c}^{\kappa}_S(m,n) - \mathrm{c}^{\kappa}_S(m,n)_+ = (-4\pi^2)m\,\mathrm{Kl}^1_S(m,n)$. To see this observe that for fixed positive K we have

$$\sum_{\llbracket\chi\rrbracket\in S_{\leq K}}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket\sum_{n>0}\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)\,\mathrm{e}(nz) = \sum_{n>0}\mathrm{Kl}^1_{S_{\leq K}}(m,n)\,\mathrm{e}(nz) \tag{3.3.22}$$

since the left hand side converges absolutely. Now $\mathrm{Kl}^1_{S_{\leq K}}(m,n) = \mathrm{Kl}^1_S(m,n)+O(n)$ as $K\to\infty$ according to (3.2.10) since the $\Sigma(x)$ there is just $\mathrm{Kl}^1_{S_{\leq x}}(m,n)$ (cf. (3.2.9)). Then (3.2.6) implies $\mathrm{Kl}^1_{S_{\leq K}}(m,n) = O(n)$ uniformly in K and n and we conclude that the limit as $K\to\infty$ of the right hand side of (3.3.22) is $\sum_{n>0}\mathrm{Kl}^1_S(m,n)\,\mathrm{e}(nz)$, and thus

$$R_0 = \mathrm{c}^{\kappa}_S(m,0)/2+\sum_{n>0}\mathrm{c}^{\kappa}_S(m,n)_0\,\mathrm{e}(nz) \tag{3.3.23}$$

as was claimed. Taking (3.3.23) together with (3.3.17) we arrive at the required identity (3.3.11) in case $\kappa=0$. The proof is complete. □

The next result identifies the Fourier series expansion of the conjugate Rademacher sum $\bar{R}^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$. As such it is a natural counterpart to Theorem 3.3.2. The method of proof is directly analogous to that of Theorem 3.3.2.

Theorem 3.3.3. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$*, and let* $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ *be cusps of* Γ*. Let* $\kappa, m\in\mathbb{Z}$ *such that* $\kappa\leq 0$ *and* $m<0$*. Then the limit defining the conjugate Rademacher sum* $\bar{R}^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ *converges locally uniformly to an anti-holomorphic function on* $\mathbb{H}$*, and we have*

$$\bar{R}^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)=\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(m\bar{z})-\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0)-F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(\bar{z})_{\mathrm{sing}}. \tag{3.3.24}$$

Theorems 3.3.2 and 3.3.3 furnish the Fourier expansions of the classical and conjugate Rademacher sums of non-positive weight. Directly analogous methods can be used to determine expressions for the Fourier coefficients of the Rademacher sums of positive weight, which are, after all, just the holomorphic Poincaré series (cf. §3.1). Indeed, there are fewer technical difficulties in the case of positive weight, since the sum appearing in the definition (3.1.2) of the Poincaré series $P^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}$ is absolutely and locally uniformly convergent (at least when $\kappa>1$). For completeness we state an analogue of Theorem 3.3.2 for $\kappa>0$. The result is standard (cf. [Iwa02]).

Theorem 3.3.4. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$*, and let* $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ *be cusps of* Γ*. Let* $\kappa,m\in\mathbb{Z}$ *such that* $\kappa>0$ *and* $m\geq 0$*. Then for* $z\in\mathbb{H}$ *we have*

$$R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)=\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz)+F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}. \tag{3.3.25}$$

Recall from §3.1 that we write $R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for $R^{0(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ and interpret $\bar{R}^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ similarly. To conclude this section we consider the difference $R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)-\bar{R}^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$. Combining Theorems 3.3.2 and 3.3.3 we obtain the following identification of the Fourier expansion of this function in terms of the exponential function $\mathrm{e}(z)$ and its conjugate $\mathrm{e}(-\bar{z})$.

Theorem 3.3.5. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$*, let* $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ *be cusps of* Γ*, and let* $m<0$*. Then we have*

$$\begin{aligned}R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)-\bar{R}^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)&=\delta_{\Gamma,\mathsf{p}|\mathsf{q}}(\mathrm{e}(mz)-\mathrm{e}(m\bar{z}))\\&\quad+F^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(\bar{z})_{\mathrm{sing}}+\mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0)+F^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}\end{aligned} \tag{3.3.26}$$

for $z\in\mathbb{H}$*.*

3.4 Variance

Recall from §3.1 that we write $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ for $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ in case q is the infinite cusp $\Gamma\cdot\infty$. Suppose that Γ is commensurable with the modular group and has width one at infinity (cf. §2.6). We will establish in this section that the Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$, once corrected by the addition of a certain constant function, is an automorphic integral of weight 2κ for Γ whenever $\kappa=1$

and $m > 0$ or $\kappa > 1$ and $m \geq 0$ or $\kappa \leq 0$ and $m < 0$. This is quite general since, according to the discussion of §3.1, any Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ can be expressed in the form $R^{\kappa(m)}_{\Gamma',\mathsf{p}'}(z)$ for some group Γ' having width one at infinity (cf. (3.1.12)).

Let us begin by considering the weight 2κ action of $G(\mathbb{Q})$ on the Rademacher component function $R^{\kappa(m)}_{\llbracket\chi}(z)$ of (3.1.4).

Lemma 3.4.1. *Let* $\llbracket\chi \in \llbracket G(\mathbb{Q})^{\times}$ *and* $\sigma \in G(\mathbb{Q})$, *and let* $\kappa, m \in \mathbb{Z}$. *In case* $\kappa > 0$ *and* $m \geq 0$ *we have*

$$\left(R^{\kappa(m)}_{\llbracket\chi} \Big|^{\kappa} \sigma \right)(z) = R^{\kappa(m)}_{\llbracket\chi\sigma}(z), \tag{3.4.1}$$

and in case $\kappa \leq 0$ *and* $m < 0$ *we have*

$$\left(R^{\kappa(m)}_{\llbracket\chi} \Big|^{\kappa} \sigma \right)(z) = R^{\kappa(m)}_{\llbracket\chi\sigma}(z) + m^{1-2\kappa} \left(J^{\kappa}_{\infty\cdot\llbracket\sigma} P^{1-\kappa(-m)}_{\llbracket\chi\sigma} \right)(z). \tag{3.4.2}$$

Proof. Note that we have

$$\left(P^{\kappa(m)}_{\llbracket\chi} |^{\kappa} \sigma \right)(z) = P^{\kappa(m)}_{\llbracket\chi\sigma}(z) \tag{3.4.3}$$

for all $\kappa, m \in \mathbb{Z}$, where $P^{\kappa(m)}_{\llbracket\chi}(z) = \mathrm{e}(m \llbracket\chi \cdot z)\,\mathrm{jac}(\llbracket\chi\,, z)^{\kappa}$ (cf. (3.1.1)), so the identity (3.4.1) follows from (3.4.3) and the fact that $R^{\kappa(m)}_{\llbracket\chi}(z) = P^{\kappa(m)}_{\llbracket\chi}(z)$ in case $\kappa > 0$, by the definition of $R^{\kappa(m)}_{\llbracket\chi}(z)$ (cf. §3.1). For (3.4.2) we suppose $\kappa \leq 0$ and $m < 0$ and use Lemma 3.1.1 to write the Rademacher component function $R^{\kappa(m)}_{\llbracket\chi}(z)$ in terms of the functions $P^{\kappa(m)}_{\llbracket\chi}(z)$ and the integral operator J^{κ}_{w}. Then, employing (3.4.3) and the identity (2.7.6), we compute

$$\begin{aligned}\left(R^{\kappa(m)}_{\llbracket\chi} \Big|^{\kappa} \sigma \right)(z) &= \left(P^{\kappa(m)}_{\llbracket\chi} \Big|^{\kappa} \sigma \right)(z) - m^{1-2\kappa} \left(\left(J^{\kappa}_{\infty\cdot\llbracket\chi} P^{1-\kappa(-m)}_{\llbracket\chi} \right) \Big|^{\kappa} \sigma \right)(z) \\ &= P^{\kappa(m)}_{\llbracket\chi\sigma}(z) - m^{1-2\kappa} \left(J^{\kappa}_{\infty\cdot\llbracket\chi\sigma} P^{1-\kappa(-m)}_{\llbracket\chi\sigma} \right)(z) + m^{1-2\kappa} \left(J^{\kappa}_{\infty\cdot\llbracket\sigma} P^{1-\kappa(-m)}_{\llbracket\chi\sigma} \right)(z).\end{aligned} \tag{3.4.4}$$

Applying Lemma 3.1.1 once again we recognize the first two terms on the right hand side (of the second line) of (3.4.4) to be $R^{\kappa(m)}_{\llbracket\chi\sigma}(z)$. We thus obtain the required identity (3.4.2). □

As for technical tools, in addition to Lemma 3.3.1, we also use the following result which shows that there is some flexibility in the convergence of the limit (3.1.6) defining the classical Rademacher sum.

Lemma 3.4.2. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$, *let* $\mathsf{p} \in \mathsf{P}_{\Gamma}$ *be a cusp of* Γ *and let* $\sigma \in G(\mathbb{Q})$. *Let* $\kappa, m \in \mathbb{Z}$ *such that either* $\kappa = 1$ *and* $m > 0$ *or* $\kappa > 1$ *and* $m \geq 0$ *or* $\kappa \leq 0$ *and*

$m < 0$. Then we have

$$\lim_{K\to\infty} \sum_{\llbracket\chi\in(U_{\leq K})\sigma} R^{\kappa(m)}_{\llbracket\chi}(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in(U\sigma)_{\leq K}} R^{\kappa(m)}_{\llbracket\chi}(z) \tag{3.4.5}$$

for $U = \llbracket\Sigma_{\mathsf{p}}^{-1}$ and $z \in \mathbb{H}$.

Proof. For $\kappa > 1$ and $m \geq 0$ both sums in (3.4.5) are absolutely convergent, and the identity certainly holds. In case $\kappa = 1$ and $m > 0$, or $\kappa \leq 0$ and $m < 0$, the identity (3.4.5) is established via the argument of Lemma 4.3 in [Nie74]. Note that this argument fails in the case that $\kappa = 1$ and $m = 0$, and indeed, the identity (3.4.5) does not hold in general when $\kappa = 1$ and $m = 0$. ☐

Theorem 3.4.3. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa > 1$ and $m \geq 0$ or $\kappa = 1$ and $m > 0$. Then the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is a modular form of weight 2κ for Γ, and is a cusp form in case $m > 0$. Further, for $\mathsf{q} \in \mathsf{P}_\Gamma$ another cusp of Γ the Fourier expansion of the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is the expansion of $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ at the cusp q.*

Proof. Let κ and m be as in the statement of the proposition. Since Γ has width one at infinity we may take Γ to be the scaling coset for Γ at the infinite cusp. Let Σ_{p} be a scaling coset for Γ at p and set $U = \llbracket\Sigma_{\mathsf{p}}^{-1}$. Then in the case that $\kappa > 1$ we have

$$R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z) = \sum_{\llbracket\chi\in U} R^{\kappa(m)}_{\llbracket\chi}(z) = \sum_{\llbracket\chi\in U} \mathrm{e}(m\,\llbracket\chi\cdot z)\,\mathrm{jac}(\llbracket\chi\,,z)^{\kappa}, \tag{3.4.6}$$

with the sum(s) in (3.4.6) converging absolutely for $z \in \mathbb{H}$. Let $\sigma \in G(\mathbb{Q})$. Then we have

$$\left(R^{\kappa(m)}_{\Gamma,\mathsf{p}}\Big|^{\kappa}\sigma\right)(z) = \sum_{\llbracket\chi\in U} \left(R^{\kappa(m)}_{\llbracket\chi}\Big|^{\kappa}\sigma\right)(z) = \sum_{\llbracket\chi\in U} R^{\kappa(m)}_{\llbracket\chi\sigma}(z) = \sum_{\llbracket\chi\in U\sigma} R^{\kappa(m)}_{\llbracket\chi}(z) \tag{3.4.7}$$

by Lemma 3.4.1, with all sums absolutely convergent, locally uniformly in z. Taking $\sigma \in \Gamma$ we see that $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is an unrestricted modular form of weight 2κ for Γ since $\Sigma_{\mathsf{p}}^{-1}\sigma = \Sigma_{\mathsf{p}}^{-1}$ for $\sigma \in \Gamma$. Taking $\mathsf{q} \in \mathsf{P}_\Gamma$ and $\sigma \in \Sigma_{\mathsf{q}}$ we see that Theorem 3.3.4 gives the Fourier expansion of $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ at q, in the sense of §2.7, since $\Sigma_{\mathsf{p}}^{-1}\sigma = \Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ for any $\sigma \in \Sigma_{\mathsf{q}}$. We conclude that the (a priori unrestricted) modular form $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is vanishing at every cusp $\mathsf{q} \in \mathsf{P}_\Gamma$, except possibly for $\mathsf{q} = \mathsf{p}$, and vanishes even there just when $m > 0$, and thus the claim is verified for $\kappa > 1$.

Consider now the case that $\kappa = 1$. We have

$$R^{1(m)}_{\Gamma,\mathsf{p}}(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} R^{1(m)}_{\llbracket\chi}(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} \mathrm{e}(m\,\llbracket\chi\cdot z)\,\mathrm{jac}(\llbracket\chi\,,z), \tag{3.4.8}$$

and it is necessary to consider (some kind of) a limit since the convergence is not absolute in this case. Let $\sigma \in G(\mathbb{Q})$ and consider the weight 2 action of σ on $R^{1(m)}_{\Gamma,\mathsf{p}}(z)$.

$$\left(R^{1(m)}_{\Gamma,\mathsf{p}}\Big|_1 \sigma\right)(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} \left(R^{1(m)}_{\llbracket\chi}\Big|_1 \sigma\right)(z) = \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} R^{1(m)}_{\llbracket\chi\sigma}(z). \tag{3.4.9}$$

Taking $\sigma \in \Gamma$ and applying Lemma 3.4.2 we see that $R^{1(m)}_{\Gamma,\mathsf{p}}(z)$ is a(n unrestricted) modular form for Γ. Applying Lemma 3.4.2 in the case that $\sigma \in \Sigma_{\mathsf{q}}$ for some cusp $\mathsf{q} \in \mathsf{P}_\Gamma$, we see that the Fourier expansion of the function $R^{1(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is the expansion of $R^{1(m)}_{\Gamma,\mathsf{p}}(z)$ at q in the sense of §2.7, so Theorem 3.3.4 applies also when $\kappa = 1$, and thus $R^{1(m)}_{\Gamma,\mathsf{p}}(z)$ is a modular form of weight 2 for Γ and even a cusp form since $m > 0$. This completes the proof. ☐

We now present an analogue of Theorem 3.4.3 for the case that $\kappa \leq 0$.

Theorem 3.4.4. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z) + \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)/2$ is an automorphic integral of weight 2κ for Γ, and for $\mathsf{q} \in \mathsf{P}_\Gamma$ another cusp of Γ the Fourier expansion of the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) + \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0)/2$ is the expansion of $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z) + \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)/2$ at the cusp q.*

Proof. Let κ and m be as in the statement of the theorem. Since Γ has width one at infinity we take Γ to be the scaling coset for Γ at the infinite cusp. Let Σ_{p} be a scaling coset for Γ at p and set $U = \llbracket\Sigma_{\mathsf{p}}^{-1}$. Let $\sigma \in G(\mathbb{Q})$. Using Lemma 3.4.1 we compute

$$\begin{aligned}\left(R^{\kappa(m)}_U\Big|^{\kappa}\sigma\right)(z) &= \lim_{K\to\infty} \sum_{\llbracket\chi\in U_{\leq K}} R^{\kappa(m)}_{\llbracket\chi\sigma}(z) + m^{1-2\kappa}\left(J^{\kappa}_{\infty\cdot\llbracket\sigma} P^{1-\kappa(-m)}_{\llbracket\chi\sigma}\right)(z)\\ &= \lim_{K\to\infty} \sum_{\llbracket\chi\in (U_{\leq K})\sigma} R^{\kappa(m)}_{\llbracket\chi}(z) + m^{1-2\kappa}\left(J^{\kappa}_{\infty\cdot\llbracket\sigma} P^{1-\kappa(-m)}_{\llbracket\chi}\right)(z)\end{aligned} \tag{3.4.10}$$

for the weight 2κ action of σ on $R^{\kappa(m)}_U(z) = R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$. Now Lemma 3.4.2 shows that

$$\left(R^{\kappa(m)}_U\Big|^{\kappa}\sigma\right)(z) = R^{\kappa(m)}_{U\sigma}(z) + m^{1-2\kappa} \lim_{K\to\infty} \sum_{\llbracket\chi\in (U_{\leq K})\sigma} \left(J^{\kappa}_{\infty\cdot\llbracket\sigma} P^{1-\kappa(-m)}_{\llbracket\chi}\right)(z), \tag{3.4.11}$$

suggesting that $R^{\kappa(m)}_U(z)$ might be an automorphic integral of weight 2κ for Γ with associated cusp form $-m^{1-2\kappa}P^{1-\kappa(-m)}_U(z)$ (cf. §2.7). This is not accurate, however, since the limit in (3.4.11) does not in general commute with the integral operator $J^{\kappa}_{\infty\cdot\llbracket\sigma}$. Rather, by the argument of Lemma 4.4. in [Nie74], we have the identity

$$\begin{aligned}\lim_{K\to\infty}\left(J^{\kappa}_{\infty\cdot[\![\sigma}\left(\sum_{[\![\chi\in(U_{\leq K})\sigma}P^{1-\kappa(-m)}_{[\![\chi}\right)\right)(z)-\left(J^{\kappa}_{\infty\cdot[\![\sigma}\left(\lim_{K\to\infty}\sum_{[\![\chi\in(U_{\leq K})\sigma}P^{1-\kappa(-m)}_{[\![\chi}\right)\right)(z)\\ =-\frac{m^{2\kappa-1}}{2}\left(\mathrm{c}^{\kappa}_{S}(m,0)\,\mathrm{jac}(\sigma,z)^{\kappa}-\mathrm{c}^{\kappa}_{S'}(m,0)\right)\end{aligned} \tag{3.4.12}$$

where $S=\{[\![\chi]\!]\mid[\![\chi\in U^{\times}\}$ and $S'=\{[\![\chi]\!]\mid[\![\chi\in(U\sigma)^{\times}\}$. Lemma 3.3.1 plays a crucial rôle in the verification of (3.4.12). Applying (3.4.12) to (3.4.11) we find that

$$\left(R^{\kappa(m)}_{U}+\frac{1}{2}\mathrm{c}^{\kappa}_{S}(m,0)\right)\Big|^{\kappa}\sigma=R^{\kappa(m)}_{U'}+\frac{1}{2}\mathrm{c}^{\kappa}_{S'}(m,0)+m^{1-2\kappa}J^{\kappa}_{\infty\cdot[\![\sigma}P^{1-\kappa(-m)}_{U'} \tag{3.4.13}$$

for $U'=U\sigma$. Taking $\sigma\in\Gamma$ and applying the identity (4.4.29) we obtain

$$\left(R^{\kappa(m)}_{\Gamma,\mathsf{p}}+\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)\right)\Big|^{\kappa}\sigma=R^{\kappa(m)}_{\Gamma,\mathsf{p}}+\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)+m^{1-2\kappa}J^{\kappa}_{\infty\cdot[\![\sigma}P^{1-\kappa(-m)}_{\Gamma,\mathsf{p}} \tag{3.4.14}$$

from (3.4.13), demonstrating that the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)+\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)/2$ is an unrestricted automorphic integral of weight 2κ for Γ with associated cusp form $-m^{1-2\kappa}P^{1-\kappa(-m)}_{\Gamma,\mathsf{p}}(z)$. Letting $\mathsf{q}\in\mathsf{P}_{\Gamma}$ and applying (3.4.13) with $\sigma\in\Sigma_{\mathsf{q}}$ we obtain

$$\begin{aligned}R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}+\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0)&=\left(R^{\kappa(m)}_{\Gamma,\mathsf{p}}+\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)\right)\Big|^{\kappa}\sigma-m^{1-2\kappa}J^{\kappa}_{\infty\cdot[\![\sigma}P^{1-\kappa(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}\\&=\left(R^{\kappa(m)}_{\Gamma,\mathsf{p}}+\frac{1}{2}\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)+m^{1-2\kappa}J^{\kappa}_{\sigma\cdot\infty}P^{1-\kappa(-m)}_{\Gamma,\mathsf{p}}\right)\Big|^{\kappa}\sigma,\end{aligned} \tag{3.4.15}$$

demonstrating that the Fourier expansion of the function $R^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)+\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0)/2$ is the expansion at q of the automorphic integral $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)+\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)/2$ in the sense of §2.7. Now Theorem 3.3.2 confirms that $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)+\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}}(m,0)/2$ is meromorphic at the cusps of Γ, and is thus an automorphic integral of weight 2κ for Γ. This completes the proof of the proposition. □

We now specialize to the case that $\kappa=0$. Our final goal in this section is to establish the Γ-invariance of the harmonic function $R^{(m)}_{\Gamma,\mathsf{p}}(z)-\bar{R}^{(m)}_{\Gamma,\mathsf{p}}(z)$ for $m<0$ in case Γ has width one at infinity. For this we utilize the spectral theory of the hyperbolic Laplacian $\Delta=(z-\bar{z})^{2}\partial_{z}\partial_{\bar{z}}$.

Consider the *Whittaker functions* $V(z,s)$ and $W(z,s)$, defined for $z,s\in\mathbb{C}$ with $z\notin\mathbb{R}$ by setting

$$V(z,s) = \mathrm{e}(z)\Gamma(s)\Phi(s,2s,z-\bar{z})(2\pi\mathbf{i}(\bar{z}-z))^s, \tag{3.4.16}$$

$$W(z,s) = \frac{\sin(\pi s)}{\sin(2\pi s)}\left(V(z,s) - V(z,1-s)\right), \tag{3.4.17}$$

in case $\Im(z) > 0$ where the function Φ is defined in (2.1.3). The values $V(z,s)$ for $\Im(z) < 0$ are obtained by imposing the symmetry $V(z,s) = V(\bar{z},s)$ and similarly for $W(z,s)$. Then $V(z,s)$ and $W(z,s)$ are eigenfunctions for Δ with eigenvalue $s(1-s)$, and the identity $V(z,1) = \mathrm{e}(\bar{z}) - \mathrm{e}(z)$ hints at a connection with the right hand side of (3.3.26). Define the *free space Green's function* $G(w,z,s)$, for $w,z \in \mathbb{H}$ with $w \neq z$ and $s \in \mathbb{C}$ such that $\Re(s) > 1$, by setting

$$G(w,z,s) = \frac{\Gamma(s)^2}{\Gamma(2s)}\,{}_2F_1\left(s,s;2s;h\right)h^s = \sum_{k\geq 0}\frac{\Gamma(k+s)^2}{\Gamma(k+2s)\Gamma(k+1)}h^{k+s}, \tag{3.4.18}$$

where ${}_2F_1(a,b;c;x)$ denotes the Gauss hypergeometric function and $h = h(w,z)$ is given as follows and depends only on the hyperbolic distance between w and z.

$$h(w,z) = \frac{(w-\bar{w})(z-\bar{z})}{(w-z)(\bar{w}-\bar{z})} = \frac{4\Im(w)\Im(z)}{|w-z|^2} \tag{3.4.19}$$

Then the function $z \mapsto G(w,z,s)$ satisfies $\Delta G(w,z,s) = s(1-s)G(w,z,s)$ and the integral operator with kernel $G(w,z,s)$ furnishes a right-inverse to the differential operator $f \mapsto (\Delta - \lambda)f$ when $\lambda = s(1-s)$.

Let Γ be a group commensurable with $G(\mathbb{Z})$, let p and q be cusps of Γ, and let Σ_{p} and Σ_{q} be scaling cosets for Γ at p and q, respectively. Define $G_{\Gamma,\mathsf{p}|\mathsf{q}}(w,z,s)$ for $w,z \in \mathbb{H}$ such that $\Gamma \cdot w \neq \Gamma \cdot z$ by setting

$$G_{\Gamma,\mathsf{p}|\mathsf{q}}(w,z,s) = \sum_{\chi \in \Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} G(w,\chi\cdot z,s). \tag{3.4.20}$$

The series (3.4.20) converges absolutely and locally uniformly when $\Re(s) > 1$ and the functions $w \mapsto G_{\Gamma,\mathsf{p}|\mathsf{q}}(w,z,s)$ and $z \mapsto G_{\Gamma,\mathsf{p}|\mathsf{q}}(w,z,s)$ are invariant for the actions of Γ^{p} and Γ^{q}, respectively. According to [Hej83] (see also [Iwa02]) we have the Fourier–Whittaker series expansion

$$\begin{split} G_{\Gamma,\mathsf{p}|\mathsf{q}}(w,z,s) = \\ \sum_{m\in\mathbb{Z}^+} G^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)W(-m\bar{w},s) + G^{(0)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)\Im(w)^{1-s} + \sum_{m\in\mathbb{Z}^+} G^{(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)W(mw,s), \end{split} \tag{3.4.21}$$

converging absolutely and locally uniformly when $\Re(s) > 1$ so long as $\Im(w) > \Im(z)$ and $\Im(z)\Im(w) > \varrho_{\Gamma,\mathsf{p}|\mathsf{q}}$, for a certain constant $\varrho_{\Gamma,\mathsf{p}|\mathsf{q}}$ depending only on Γ and p and q, where the coefficient functions $G^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ are Γ^{q}-invariant, and themselves admit the following series expansions of Poincaré type, converging absolutely for $\Re(s) > 1$ and locally uniformly for $\Im(z) > 0$.

$$G^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \frac{1}{m} \sum_{[\![\chi \in [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} V(m\,[\![\chi \cdot z, s) \tag{3.4.22}$$

$$G^{(0)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \frac{4\pi}{2s-1} \sum_{[\![\chi \in [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} \Im([\![\chi \cdot z)^{s} \tag{3.4.23}$$

$$G^{(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \frac{1}{m} \sum_{[\![\chi \in [\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} V(-m\,[\![\chi \cdot \bar{z}, s) \tag{3.4.24}$$

Let $m > 0$. Then we have the following analogue of (3.4.21) for the Green's function coefficient $G^{(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ which is again reminiscent of the right hand side of (3.3.26).

$$\begin{gathered} G^{(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\frac{1}{m}V(-m\bar{z},s) + \\ \sum_{n>0} \mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,-n,s)W(-n\bar{z},s) + \mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,0,s)\Im(z)^{1-s} + \sum_{n>0} \mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,n,s)W(nz,s) \end{gathered} \tag{3.4.25}$$

The coefficients in (3.4.25) may be defined via the following formulas, in which we assume $m, n > 0$.

$$\begin{gathered} \mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,-n,s) = \\ \frac{1}{m^{1-s}n^{s}} \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \sum_{k\geq 0} \mathrm{e}(-m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}(-n\,[\![\chi^{-1}]\!]\cdot\infty)(4\pi^2\varrho\,[\![\chi]\!])^{k+s} m^{(k)} n^{(k+2s-1)} \end{gathered} \tag{3.4.26}$$

$$\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,0,s) = \frac{1}{m^{1-s}}\frac{4\pi}{2s-1} \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \sum_{k\geq 0} \mathrm{e}(-m\,[\![\chi]\!]\cdot\infty)\varrho\,[\![\chi]\!]^{s}\,\frac{\pi^{s}}{\Gamma(s)} \tag{3.4.27}$$

$$\begin{gathered} \mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(-m,n,s) = \\ \frac{1}{m^{1-s}n^{s}} \lim_{K\to\infty} \sum_{[\![\chi]\!]\in S_{\leq K}} \sum_{k\geq 0} \mathrm{e}(-m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}(n\,[\![\chi^{-1}]\!]\cdot\infty)(-1)^{k}(4\pi^2\varrho\,[\![\chi]\!])^{k+s} m^{(k)} n^{(k+2s-1)} \end{gathered} \tag{3.4.28}$$

Comparing with Lemma (3.2.6) we obtain a direct relationship between the coefficient functions $\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n,s)$ arising from the Green's function, and the coefficient functions $\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n)$ of §3.2 which arise from the Rademacher sums.

Proposition 3.4.5. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\mathsf{p},\mathsf{q} \in \mathsf{P}_{\Gamma}$, and let $\kappa > 0$*

and $m < 0$. *Then we have the following identities for* $n > 0$.

$$\mathrm{c}^{1-\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,-n) = -m^{\kappa}n^{1-\kappa}\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,-n,\kappa) \tag{3.4.29}$$

$$\mathrm{c}^{1-\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0) = -m^{\kappa}\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,0,\kappa) \tag{3.4.30}$$

$$\mathrm{c}^{1-\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n) = -m^{\kappa}n^{1-\kappa}\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n,\kappa) \tag{3.4.31}$$

Taking $\kappa = 1$ in Proposition 3.4.5 we obtain the identity $\mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n) = -m\mathrm{g}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n,1)$ for $m < 0$, which, by Theorem 3.3.5, the series expansion (3.4.25) and the identity $W(z,1) = \mathrm{e}(z)$, may be reformulated as

$$R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) - \bar{R}^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = -mG^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,1) \tag{3.4.32}$$

for $m < 0$. The Γ^{q}-invariance of the difference $R^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) - \bar{R}^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ of classical and conjugate Rademacher sums of weight 0 now follows from the Γ^{q}-invariance of the Green's function coefficient $G^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ at $s = 1$ (cf. [Nie73]). In particular, we have the following result.

Theorem 3.4.6. *Let* Γ *be a group commensurable with* $G(\mathbb{Z})$ *that has width one at infinity, let* $\mathsf{p} \in \mathsf{P}_{\Gamma}$ *be a cusp of* Γ, *and let* $m < 0$. *Then the harmonic function* $R^{(m)}_{\Gamma,\mathsf{p}}(z) - \bar{R}^{(m)}_{\Gamma,\mathsf{p}}(z)$ *is* Γ*-invariant.*

4 Modified Rademacher sums

We have seen in §3.4 that the classical Rademacher sum $R^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ of §3 is an automorphic integral for Γ only after the addition of a particular constant function, which does not generally vanish. In this section we introduce a modification of Rademacher's construction via which the correct constant term appears naturally. Our approach employs an analytic continuation of the component functions defining the classical Rademacher sums, and entails the assignment of a Dirichlet series to each triple $(\Gamma,\mathsf{p},\mathsf{q})$ where Γ is a group commensurable with the modular group and $\mathsf{p},\mathsf{q} \in \mathsf{P}_{\Gamma}$ are cusps for Γ.

4.1 Construction

For $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$, and for $\llbracket\chi \in \llbracket G(\mathbb{Q})^{\times}$ we define the *continued Rademacher component function*, denoted $(z,s) \mapsto T^{\kappa(m)}_{\llbracket\chi}(z,s)$, by setting

$$T^{\kappa(m)}_{\llbracket\chi}(z,s) = \mathrm{e}(m\llbracket\chi\cdot z)\,\mathrm{t}^{\kappa}(m,\llbracket\chi,z,s)\,\mathrm{jac}(\llbracket\chi,z)^{\kappa} \tag{4.1.1}$$

for $z \in \mathbb{H}$ and $s \in \mathbb{C}$ such that $\Re(s) > 1$ where $\mathrm{t}^{\kappa}(m, [\![\chi\,, z, s)$ is the *continued Rademacher regularization factor of weight* 2κ which is in turn given by

$$\mathrm{t}^{\kappa}(m, [\![\chi\,, z, s) = \Phi(s - 2\kappa, 1 + s - 2\kappa, m\,[\![\chi \cdot \infty - m\,[\![\chi \cdot z)(2\pi\mathbf{i}(m\,[\![\chi \cdot z - m\,[\![\chi \cdot \infty))^{s-2\kappa} \tag{4.1.2}$$

in case $[\![\chi \in [\![G(\mathbb{Q})^{\times}$ and $\mathrm{t}^{\kappa}(m, [\![\chi\,, z, s) = 1$ otherwise. We also define a function $s \mapsto T^{\kappa(m)}_{0,[\![\chi}(s)$ by setting

$$\begin{aligned} T^{\kappa(m)}_{0,[\![\chi}(s) =\, & \mathrm{e}(m\,[\![\chi \cdot \infty)\,(2\pi\mathbf{i}(m\,[\![\chi \cdot 0 - m\,[\![\chi \cdot \infty))^{(s-2\kappa)}\,\mathrm{jac}([\![\chi\,, 0)^{\kappa} \\ & - \mathrm{e}(s/2)\,\mathrm{e}(m\,[\![\chi \cdot \infty)\,(2\pi\mathbf{i}(m\,[\![\chi \cdot \infty - m\,[\![\chi \cdot 0))^{(s-2\kappa)}\,\mathrm{jac}([\![\chi\,, 0)^{\kappa} \end{aligned} \tag{4.1.3}$$

in case $[\![\chi \in [\![G(\mathbb{Q})^{\times\times}$, and by setting $T^{\kappa(m)}_{0,[\![\chi}(s) = 0$ otherwise, and we define the *modified continued Rademacher component function*, denoted $(z, s) \mapsto Q^{\kappa(m)}_{[\![\chi}(z, s)$, by subtracting $T^{\kappa(m)}_{0,[\![\chi}(z, s)$ from $T^{\kappa(m)}_{[\![\chi}(z, s)$.

$$Q^{\kappa(m)}_{[\![\chi}(z, s) = T^{\kappa(m)}_{[\![\chi}(z, s) - T^{\kappa(m)}_{0,[\![\chi}(z, s) \tag{4.1.4}$$

Now for $U \subset [\![G(\mathbb{Q})$, and for $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$, we define the *continued Rademacher sum* $T^{\kappa(m)}_{U}(z, s)$, and the *modified continued Rademacher sum* $Q^{\kappa(m)}_{U}(z, s)$, by setting

$$T^{\kappa(m)}_{U}(z, s) = \sum_{[\![\chi \in U} T^{\kappa(m)}_{[\![\chi}(z, s), \tag{4.1.5}$$

$$Q^{\kappa(m)}_{U}(z, s) = \sum_{[\![\chi \in U} Q^{\kappa(m)}_{[\![\chi}(z, s), \tag{4.1.6}$$

and we define $T^{\kappa(m)}_{U}(z)$ and $Q^{\kappa(m)}_{U}(z)$ by taking the limit as s tends to 1 through the region $\Re(s) > 1$ in $T^{\kappa(m)}_{U}(z, s)$ and $Q^{\kappa(m)}_{U}(z, s)$, respectively, so long as these limits exist.

$$T^{\kappa(m)}_{U}(z) = \lim_{s\to 1^+} T^{\kappa(m)}_{U}(z, s) \tag{4.1.7}$$

$$Q^{\kappa(m)}_{U}(z) = \lim_{s\to 1^+} Q^{\kappa(m)}_{U}(z, s) \tag{4.1.8}$$

The function $Q^{\kappa(m)}_{U}(z)$ is the most important from the point of view of automorphy for subgroups of $G(\mathbb{Q})$, while the function $T^{\kappa(m)}_{U}(z)$ plays a special rôle in the critical case that $\kappa = 0$. We call $Q^{\kappa(m)}_{U}(z)$ the *modified Rademacher sum of weight* 2κ *and order* m *associated to* U, and we call $T^{\kappa(m)}_{U}(z)$ the *normalized Rademacher sum of weight* 2κ *and order* m *associated to* U.

Observe that we recover the Rademacher component function $R^{\kappa(m)}_{[\![\chi}(z)$ by taking $s = 1$

in either the continued Rademacher component function $T^{\kappa(m)}_{\llbracket\chi}(z,s)$ or the modified continued Rademacher component function $Q^{\kappa(m)}_{\llbracket\chi}(z,s)$.

Lemma 4.1.1. *Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$, and let $\llbracket\chi \in \llbracket G(\mathbb{Q})$. Then we have $T^{\kappa(m)}_{\llbracket\chi}(z,1) = R^{\kappa(m)}_{\llbracket\chi}(z)$ and $T^{\kappa(m)}_{0,\llbracket\chi}(1) = 0$, and hence $Q^{\kappa(m)}_{\llbracket\chi}(z,1) = R^{\kappa(m)}_{\llbracket\chi}(z)$.*

Despite the result of Lemma 4.1.1, it is generally not the case that the modified Rademacher sum $Q^{\kappa(m)}_{U}(z)$ and the classical Rademacher sum $R^{\kappa(m)}_{U}(z)$ coincide (cf. Proposition 4.4.10).

We typically take U to be of the form $U = \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ where $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ is a system of scaling cosets (cf. §2.6) for some group Γ commensurable with the modular group, and $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ are cusps of Γ. In this case we write $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for $Q^{\kappa(m)}_{U}(z)$ and $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for $T^{\kappa(m)}_{U}(z)$, suppressing the choice of scaling cosets from notation. A change in the choice of scaling cosets Σ_{p} and Σ_{q} replaces $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ with a function of the form $\mathrm{e}(\alpha)Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z+\beta)$ for some $\alpha, \beta \in \mathbb{Q}$, and similarly for $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$.

In the case that p or q is the *infinite cusp* $\Gamma\cdot\infty$ we omit it from notation, writing $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ for $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\Gamma\cdot\infty}(z)$, and $Q^{\kappa(m)}_{\Gamma|\mathsf{q}}(z)$ for $Q^{\kappa(m)}_{\Gamma,\Gamma\cdot\infty|\mathsf{q}}(z)$, and similarly for the functions $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$. The functions $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ and $T^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ are the most important, for we shall see in §4.5 that $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is an automorphic integral of weight 2κ for Γ with a single pole at the cusp p in case Γ has width one at infinity, and we will see in §4.4 that $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ and $T^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ differ only by a constant function. We call $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ the *modified Rademacher sum of weight 2κ and order m associated to Γ at the cusp* p, and we call $T^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ the *normalized Rademacher sum of weight 2κ and order m associated to Γ at the cusp* p. We shall see also in §4.5 that the Fourier expansion of the function $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is the expansion of $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ at the cusp q, in the sense of §2.7.

Just as for the classical Rademacher sums (cf. §3.1) we have the result that every modified Rademacher sum $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ may be expressed in the form $Q^{\kappa(m)}_{\Gamma',\mathsf{p}'}(z)$ for some group Γ' with width one at infinity, and some cusp p' of Γ', and similarly for the normalized Rademacher sums $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$. Precisely, we have

$$Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = Q^{\kappa(m)}_{\Gamma^{\mathsf{q}},\mathsf{p}^{\mathsf{q}}}(z) \tag{4.1.9}$$

$$T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = T^{\kappa(m)}_{\Gamma^{\mathsf{q}},\mathsf{p}^{\mathsf{q}}}(z) \tag{4.1.10}$$

subject to the understanding that the data defining the right hand sides of (4.1.9) and (4.1.10) is related to the data defining the left hand sides of (4.1.9) and (4.1.10) by $\Gamma^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{q}}$ and $\mathsf{p}^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1}\cdot\mathsf{p}$ and $\Sigma_{\mathsf{p}^{\mathsf{q}}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}$.

Our primary interest in this article is in the distinguished case that $\kappa = 0$. In order to simplify notation, and maintain consistency with the notation of §1, we write $Q^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ as a shorthand for $Q^{0(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$, and we write $T^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ as a shorthand for $T^{0(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$.

4.2 Coefficients

In order to recover explicit expressions of the Fourier coefficients of the modified Rademacher sums we employ the following generalizations of $\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n)$ and $\mathrm{c}^{\kappa}_{S}(m,n)$ for $\kappa,m,n\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0<n$. For $\kappa\leq 0$ and $m<0<n$ we define the entire function $s\mapsto \mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n,s)$ by setting

$$\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n,s)=(-1)^{\kappa}\sum_{k\geq 0}(4\pi^2)^{k-\kappa+s}\varrho\,[\![\chi]\!]^{k-2\kappa+s}\,(-m)^{(k-2\kappa+s)}n^{(k+s-1)}, \tag{4.2.1}$$

so that $\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n,1)=\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n)$. We then define the *continued coefficient function* $s\mapsto \mathrm{c}^{\kappa}_{S}(m,n,s)$ for $\kappa\leq 0$ and $m<0<n$ by setting

$$\mathrm{c}^{\kappa}_{S}(m,n,s)=\sum_{[\![\chi]\!]\in S}\mathrm{Kl}^{\kappa}_{[\![\chi]\!]}(m,n)\,\mathrm{Bl}^{\kappa}_{[\![\chi]\!]}(m,n,s). \tag{4.2.2}$$

Lemma 4.2.1. *The functions $s\mapsto \mathrm{c}^{\kappa}_{S}(m,n,s)$ admit the following series representation.*

$$\begin{aligned}&\mathrm{c}^{\kappa}_{S}(m,n,s)\\&=(-1)^{\kappa}\lim_{K\to\infty}\sum_{[\![\chi]\!]\in S_{\leq K}}\sum_{k\geq 0}\mathrm{e}(m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}(-n\,[\![\chi^{-1}]\!]\cdot\infty)(4\pi^2\varrho\,[\![\chi]\!])^{k-\kappa+s}(-m)^{(k-2\kappa+s)}n^{(k+s-1)}\end{aligned} \tag{4.2.3}$$

We have $\mathrm{c}^{\kappa}_{S}(m,n,1)=\mathrm{c}^{\kappa}_{S}(m,n)$.

We usually take S to be of the form $S=[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ where Σ_{p} and Σ_{q} are scaling cosets (cf. §2.6) for a group Γ commensurable with $G(\mathbb{Z})$ at cusps $\mathsf{p},\mathsf{q}\in\mathsf{P}_{\Gamma}$. We write $\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n,s)$ for $\mathrm{c}^{\kappa}_{S}(m,n,s)$ in the case that $S=[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$.

Utilizing the continued coefficient functions $s\mapsto \mathrm{c}^{\kappa}_{S}(m,n,s)$ we generalize the function $F^{\kappa(m)}_{S}(z)_{\mathrm{van}}$ in case $\kappa\leq 0$ and $m<0$ by setting

$$F^{\kappa(m)}_{S}(z,s)_{\mathrm{van}}=\sum_{n>0}\mathrm{c}^{\kappa}_{S}(m,n,s)\,\mathrm{e}(nz). \tag{4.2.4}$$

We write $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}}$ for $F^{\kappa(m)}_{S}(z,s)_{\mathrm{van}}$ in the case that $S=[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$ for Γ a group commensurable with $G(\mathbb{Z})$ and for $\Sigma_{\mathsf{p}},\Sigma_{\mathsf{q}}\in\Gamma\backslash G(\mathbb{Q})$ scaling cosets for Γ at cusps $\mathsf{p},\mathsf{q}\in\mathsf{P}_{\Gamma}$. We write $F^{(m)}_{S}(z,s)_{\mathrm{van}}$ as a shorthand for $F^{0(m)}_{S}(z,s)_{\mathrm{van}}$, and interpret the notation $F^{(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}}$ similarly.

We have the following analogue of Proposition 3.2.10 for the functions $F^{\kappa(m)}_{S}(z,s)_{\mathrm{van}}$ in case S is of the form $[\![\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}]\!]^{\times}$.

Proposition 4.2.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$, and let $\mathsf{p},\mathsf{q}\in\mathsf{P}_{\Gamma}$ be cusps of*

Γ. *Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the series*

$$F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}} = \sum_{n>0} \mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n,s)\,\mathrm{e}(nz) \tag{4.2.5}$$

converges absolutely and locally uniformly in z and s for $z \in \mathbb{H}$ and $\Re(s) \geq 1$. In particular, the assignment $z \mapsto F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}}$ is a holomorphic function on $\mathbb{H}$ whenever $\kappa \leq 0$ and $m < 0$ and $\Re(s) \geq 1$.

4.3 Dirichlet series

Given $[\![\chi]\!] \in [\![G(\mathbb{Q})]\!]^{\times}$, and $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$, we define a Dirichlet series $Z^{\kappa(m)}_{[\![\chi]\!]}(s)$ by setting

$$\begin{aligned} Z^{\kappa(m)}_{[\![\chi]\!]}(s) &= (-1)^{1-\kappa} \sum_{n \in \mathbb{Z}^{\times}} \mathrm{e}(m\,[\![\chi]\!] \cdot \infty)\,\mathrm{e}(-n\,[\![\chi^{-1}]\!] \cdot \infty)(4\pi^2 \varrho\,[\![\chi]\!])^{1-s-\kappa} |m|^{(1-s-2\kappa)} n^{(-s)} \\ &= (-1)^{1-\kappa}\,\mathrm{e}(m\,[\![\chi]\!] \cdot \infty)(4\pi^2 \varrho\,[\![\chi]\!])^{1-s-\kappa} |m|^{(1-s-2\kappa)} \sum_{n \in \mathbb{Z}^{\times}} \mathrm{e}(-n\,[\![\chi^{-1}]\!] \cdot \infty) n^{(-s)}. \end{aligned} \tag{4.3.1}$$

(Recall that $\mathbb{Z}^{\times}$ denotes the non-zero elements of $\mathbb{Z}$.) This series (4.3.1) converges absolutely and locally uniformly in the half plane $\Re(s) > 1$, and admits a meromorphic continuation to all of $\mathbb{C}$. Given $S \subset [\![G(\mathbb{Q})]\!]^{\times}$ we define the *zeta function of weight 2κ and order m associated to S*, to be denoted $Z^{\kappa(m)}_{S}(s)$, by setting

$$Z^{\kappa(m)}_{S}(s) = \sum_{[\![\chi]\!] \in S} Z^{\kappa(m)}_{[\![\chi]\!]}(s). \tag{4.3.2}$$

As it is defined here $Z^{\kappa(m)}_{S}$ is only a formal sum but it will be shown to converge, and admit an analytic continuation, for various choices of S in §4.4 (cf. Proposition 4.4.4, Theorem 4.4.5). In order to describe the analytic continuation of the zeta function $s \mapsto Z^{\kappa(m)}_{S}(s)$ explicitly, we formulate an identity which expresses it in terms of the Hurwitz zeta function. In preparation for this we define a function $s \mapsto \tilde{Z}^{\kappa(m)}_{[\![\chi]\!]}(s)$, for each $[\![\chi]\!] \in [\![G(\mathbb{Q})]\!]^{\times}$, by setting

$$\tilde{Z}^{\kappa(m)}_{[\![\chi]\!]}(s) = \mathrm{e}(m\,[\![\chi]\!] \cdot \infty)\varrho\,[\![\chi]\!]^{s-\kappa}\,(-2\pi\mathbf{i}m)^{(s-2\kappa)}(\mathrm{e}(s/2) - \mathrm{e}(-s/2))\zeta(1 - \alpha_{[\![\chi]\!]}, s) \tag{4.3.3}$$

where $\zeta(\alpha, s)$ is defined by $\zeta(\alpha, s) = \sum_{n \geq 0}(n+\alpha)^{-s}$ for $\Re(\alpha) > 0$ and $\Re(s) > 1$, and $\alpha_{[\![\chi]\!]} \in \mathbb{Q}$ is chosen so that

$$0 \leq \alpha_{[\![\chi]\!]} < 1, \quad \alpha_{[\![\chi]\!]} + \mathbb{Z} = -\chi^{-1} \cdot \infty + \mathbb{Z}. \tag{4.3.4}$$

We then define $\tilde{Z}^{\kappa(m)}_S(s)$ by setting $\tilde{Z}^{\kappa(m)}_S(s)=\sum_{[\![\chi]\!]\in S}\tilde{Z}^{\kappa(m)}_{[\![\chi]\!]}(s)$.

Lemma 4.3.1. *We have* $Z^{\kappa(m)}_S(1-s)=\tilde{Z}^{\kappa(m)}_S(s)$.

Proof. Let $[\![\chi]\!]\in[\![G(\mathbb{Q})]\!]^\times$ and choose $\alpha_{[\![\chi]\!]}\in\mathbb{Q}$ such that $0\leq\alpha_{[\![\chi]\!]}<1$ and $\alpha_{[\![\chi]\!]}+\mathbb{Z}=-\chi^{-1}\cdot\infty+\mathbb{Z}$ for any $\chi\in[\![\chi]\!]$. We will show that $Z^{\kappa(m)}_{[\![\chi]\!]}(s)=\tilde{Z}^{\kappa(m)}_{[\![\chi]\!]}(1-s)$. Observe that we have

$$\sum_{n\in\mathbb{Z}^\times}\mathrm{e}(-n\left[\!\left[\chi^{-1}\right]\!\right]\cdot\infty)n^{(-s)}=\frac{1}{\Gamma(1-s)}\left(F\left(-\left[\!\left[\chi^{-1}\right]\!\right]\cdot\infty,s\right)+\mathrm{e}(-s/2)F\left(\left[\!\left[\chi^{-1}\right]\!\right]\cdot\infty,s\right)\right),\tag{4.3.5}$$

where $F(\alpha,s)$ denotes the periodic zeta function, defined by $F(\alpha,s)=\sum_{n>0}\mathrm{e}(n\alpha)n^{-s}$ for $\Re(s)>1$. The *Hurwitz relation*, which may be stated in the form

$$F(\alpha,s)+\mathrm{e}(-s/2)F(-\alpha,s)=\frac{(-2\pi\mathbf{i})^s}{\Gamma(s)}\zeta(1-\alpha,1-s)\tag{4.3.6}$$

(cf. [KR01]) implies the identity

$$\sum_{n\in\mathbb{Z}^\times}\mathrm{e}(-n\left[\!\left[\chi^{-1}\right]\!\right]\cdot\infty)n^{(-s)}=\frac{(-2\pi\mathbf{i})^s}{\Gamma(1-s)\Gamma(s)}\zeta(1-\alpha_{[\![\chi]\!]},1-s),\tag{4.3.7}$$

so the expression (4.3.1) may be reformulated as

$$Z^{\kappa(m)}_{[\![\chi]\!]}(s)=(-1)^{1-\kappa}\,\mathrm{e}(m\,[\![\chi]\!]\cdot\infty)(4\pi^2\varrho\,[\![\chi]\!])^{1-s-\kappa}|m|^{(1-s-2\kappa)}\frac{(-2\pi\mathbf{i})^s}{\Gamma(1-s)\Gamma(s)}\zeta(1-\alpha_{[\![\chi]\!]},1-s).\tag{4.3.8}$$

We obtain the coincidence of the right hand side of (4.3.8) with $\tilde{Z}^{\kappa(m)}_{[\![\chi]\!]}(1-s)$ by applying the functional equation for the Gamma function, which may be expressed as

$$(\mathrm{e}(s/2)-\mathrm{e}(-s/2))\Gamma(s)\Gamma(1-s)=2\pi\mathbf{i}\tag{4.3.9}$$

This completes the proof of the claim. □

We write $Z^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(s)$ for $Z^{\kappa(m)}_S(s)$ in the case that $S=\left[\!\left[\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\right]\!\right]^\times$ for Γ a group commensurable with $G(\mathbb{Z})$ and for $\Sigma_{\mathsf{p}},\Sigma_{\mathsf{q}}\in\Gamma\backslash G(\mathbb{Q})$ scaling cosets for Γ at cusps $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$.

4.4 Convergence

Our main objective in this section is to furnish explicit expressions for the Fourier expansions of the functions $z\mapsto T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ and $z\mapsto Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$, defined by the continued Rademacher sums and modified continued Rademacher sums, respectively. This will yield the Fourier expansions of the modified Rademacher sums $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ and the normalized Rademacher sums $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$. In preparation for the derivation of these expansions we verify the absolute and locally uniform convergence of the series defining $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ and $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ in case $\Re(s) > 1$.

Proposition 4.4.1. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the continued Rademacher sum*

$$T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \sum_{\llbracket\chi \in \llbracket\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}} T^{\kappa(m)}_{\llbracket\chi}(z,s) \tag{4.4.1}$$

is absolutely convergent, locally uniformly in z and s, for $z \in \mathbb{H}$ and $\Re(s) > 1$.

Proof. We begin by observing that we have

$$T^{\kappa(m)}_{\llbracket\chi}(z,s) = \mathrm{e}(m\llbracket\chi\cdot\infty)\,\mathrm{e}(m\llbracket\chi\cdot z - m\llbracket\chi\cdot\infty, s-2\kappa)\,\mathrm{jac}(\llbracket\chi\,,z)^\kappa \tag{4.4.2}$$

in case $\llbracket\chi \in \llbracket G(\mathbb{Q})^\times$, and $T^{\kappa(m)}_{\llbracket\chi}(z,s) = \mathrm{e}(m\llbracket\chi\cdot z)\,\mathrm{jac}(\llbracket\chi\,,z)^\kappa$ otherwise, where $\mathrm{e}(z,s)$ denotes the generalized exponential function of (2.1.6). This identity (4.4.2) follows from an application of the *Kummer transformation* $\Phi(a,b,z) = \mathrm{e}(z)\Phi(b-a,b,-z)$ to the expression (4.1.1) defining $T^{\kappa(m)}_{\llbracket\chi}(z,s)$. The absolute and locally uniform convergence of (4.4.1) can be established by implementing methods present in the proofs of Proposition 3.2.8 and Theorem 3.3.2. Namely, the right hand side of (4.4.1) is very similar to the quantity R_+ defined in (3.3.15) and we can rewrite it so as to arrive at an analogue of (3.3.17). Then the coefficients of $\mathrm{e}(nz)$ in the resulting sum can be estimated just as the values $\mathrm{c}^\kappa_S(m,n)$ are in (3.2.27) for $\kappa > 1$. □

Proposition 4.4.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the sum*

$$\sum_{\llbracket\chi \in \llbracket\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{q}} T^{\kappa(m)}_{0,\llbracket\chi}(s) \tag{4.4.3}$$

is absolutely convergent, locally uniformly in s, for $\Re(s) > 1$.

Propositions 4.4.1 and 4.4.2 imply the absolute and locally uniform convergence of the expression defining the modified continued Rademacher sum $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ for $z \in \mathbb{H}$ and $\Re(s) > 1$.

Proposition 4.4.3. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ, and let Σ_p and Σ_q be scaling cosets for Γ at p and q, respectively. Let $\kappa \in \mathbb{Z}$ such that $\kappa \leq 0$,* *and let $m \in \mathbb{Z}^+$. Then the modified Rademacher sum*

$$Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \sum_{\llbracket\chi \in \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} T^{\kappa(m)}_{\llbracket\chi}(z,s) - T^{\kappa(m)}_{0,\llbracket\chi}(s) \tag{4.4.4}$$

converges absolutely, and locally uniformly in z and s, for $z \in \mathbb{H}$ and $\Re(s) > 1$.

We next seek to relate the modified Rademacher sums $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ to the functions $F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$ and $Z^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(s)$ of §4.2 and §4.3, respectively.

Proposition 4.4.4. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ, with scaling matrices Σ_{p} and Σ_{q}, respectively. Let $\kappa \in \mathbb{Z}$ such that $\kappa \leq 0$ and let $m \in \mathbb{Z}^+$. Then we have*

$$\sum_{\llbracket\chi \in \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}} T^{\kappa(m)}_{0,\llbracket\chi}(s) = -Z^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(1-s) \tag{4.4.5}$$

for $\Re(s) > 1$ where the convergence is absolute and locally uniform.

Proof. Set $U = \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ and $S = \llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\rrbracket^{\times}$. Recall that $T^{\kappa(m)}_{0,\llbracket\chi}(s) = 0$ when $\llbracket\chi \notin \llbracket G(\mathbb{Q})^{\times\times}$, and otherwise

$$\begin{aligned} T^{\kappa(m)}_{0,\llbracket\chi}(s) = {}& \mathrm{e}(m\llbracket\chi\cdot\infty)\,(2\pi\mathbf{i}(m\llbracket\chi\cdot 0 - m\llbracket\chi\cdot\infty))^{(s-2\kappa)}\,\mathrm{jac}(\llbracket\chi\,,0)^{\kappa} \\ & - \mathrm{e}(s/2)\,\mathrm{e}(m\llbracket\chi\cdot\infty)\,(2\pi\mathbf{i}(m\llbracket\chi\cdot\infty - m\llbracket\chi\cdot 0))^{(s-2\kappa)}\,\mathrm{jac}(\llbracket\chi\,,0)^{\kappa}, \end{aligned} \tag{4.4.6}$$

by the definition (cf. §4.1) of $T^{\kappa(m)}_{0,\llbracket\chi}(s)$. We apply the identities $\chi\cdot\infty - \chi\cdot 0 = \varrho(\chi)/(-\chi^{-1}\cdot\infty)$ and $\mathrm{jac}(\llbracket\chi\,,0) = \varrho(\chi)/(-\chi^{-1}\cdot\infty)^2$ (cf. §2.4) so as to write

$$T^{\kappa(m)}_{0,\llbracket\chi}(s) = \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket^{s-\kappa}(-2\pi\mathbf{i}m)^{(s-2\kappa)}((-\chi^{-1}\rrbracket\cdot\infty)^{-s} - \mathrm{e}(s/2)(\chi^{-1}\rrbracket\cdot\infty)^{-s}). \tag{4.4.7}$$

Let $\llbracket\chi\rrbracket \in \llbracket G(\mathbb{Q})\rrbracket^{\times}$. Then we have

$$\sum_{\llbracket\chi \in \llbracket\chi\rrbracket} T^{\kappa(m)}_{0,\llbracket\chi}(s) = \mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket^{s-\kappa}(-2\pi\mathbf{i}m)^{(s-2\kappa)}\sideset{}{^*}\sum_{n\in\mathbb{Z}}((n+\alpha_{\llbracket\chi\rrbracket})^{-s} - \mathrm{e}(s/2)(n-\alpha_{\llbracket\chi\rrbracket})^{-s}) \tag{4.4.8}$$

where $\alpha_{\llbracket\chi\rrbracket} \in \mathbb{Q}$ is chosen so that $0 \leq \alpha_{\llbracket\chi\rrbracket} < 1$ and $\alpha_{\llbracket\chi\rrbracket} + \mathbb{Z} = -\chi^{-1}\cdot\infty + \mathbb{Z}$, and the superscript in the summation $\sum^*_{n\in\mathbb{Z}}$ indicates to omit the term corresponding to $n = 0$ in case $\alpha_{\llbracket\chi\rrbracket} = 0$.

We compute

$$\sum_{n\in\mathbb{Z}}^{*}((n+\alpha_{\llbracket\chi\rrbracket})^{-s}-\mathrm{e}(s/2)(n-\alpha_{\llbracket\chi\rrbracket})^{-s})=(\mathrm{e}(-s/2)-\mathrm{e}(s/2))\zeta(1-\alpha_{\llbracket\chi\rrbracket},s) \tag{4.4.9}$$

where $\zeta(\alpha,s)$ denotes the Hurwitz zeta function (cf. §4.3), so we have

$$\sum_{\llbracket\chi\in U^{\times\times}}T^{\kappa(m)}_{0,\llbracket\chi}(s)=\sum_{\llbracket\chi\rrbracket\in S^{\times}}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\varrho\llbracket\chi\rrbracket^{s-\kappa}(-2\pi\mathbf{i}m)^{(s-2\kappa)}(\mathrm{e}(-s/2)-\mathrm{e}(s/2))\zeta(1-\alpha_{\llbracket\chi\rrbracket},s). \tag{4.4.10}$$

Comparing with the expression (4.3.3) defining $Z^{\kappa(m)}_{\llbracket\chi\rrbracket}(s)'$ we see that the right hand side of (4.4.10) coincides with $-Z^{\kappa(m)}_{S^{\times}}(s)'$. Now Lemma 4.3.1 yields the required identity (4.4.5). This completes the proof. ☐

Theorem 4.4.5. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p},\mathsf{q}\in\mathsf{P}_{\Gamma}$ be cusps of Γ. Let $\kappa,m\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0$. Then we have*

$$Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)=\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(-mz)+Z^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(1-s)+F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}} \tag{4.4.11}$$

for $z\in\mathbb{H}$ and $\Re(s)>1$.

Proof. Let $m<0<\kappa$, let $s\in\mathbb{C}$ with $\Re(s)>1$, set $U=\llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ and set $S=\llbracket\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}\rrbracket^{\times}$. By the defining properties of scaling cosets (cf. §2.6) and by the definition (4.1.1) of the continued Rademacher component function $T^{1-\kappa(m)}_{\llbracket\chi}(z,s)$ we have

$$\sum_{\llbracket\chi\in U}T^{1-\kappa(m)}_{\llbracket\chi}(z,s)=\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz)+\sum_{\llbracket\chi\in U^{\times}}T^{1-\kappa(m)}_{\llbracket\chi}(z,s) \tag{4.4.12}$$

so by Proposition 4.4.4, and the definition of the modified continued Rademacher component function $Q^{1-\kappa(m)}_{\llbracket\chi}(z,s)$, it suffices for us to show that

$$\sum_{\llbracket\chi\in U^{\times}}T^{1-\kappa(m)}_{\llbracket\chi}(z,s)=F^{1-\kappa(m)}_{S}(z,s)_{\mathrm{van}}. \tag{4.4.13}$$

We will verify the equality of (4.4.13) by using the Lipschitz summation formula (3.3.1) to transform the expression on the right hand side of (4.4.13) into that on the left. This is essentially the approach employed originally by Rademacher in [Rad39], and the reverse of that employed by Niebur in Lemma 4.2 of [Nie74], except that we can avoid the technical difficulties of [Rad39] and [Nie74], such as the need to employ the identity (3.3.10), by working only with the versions of the Lipschitz summation formula (viz., $\Re(s)>1$) in which both sides of the identity are absolutely and locally uniformly convergent series.

We begin by inspecting Lemma 4.2.1 to find that

$$\begin{aligned}&(-1)^{1-\kappa}\mathrm{c}_S^{1-\kappa}(m,n,s)\\&=\sum_{\llbracket\chi\rrbracket\in S}\sum_{k\geq 0}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\,\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)(4\pi^2\varrho\llbracket\chi\rrbracket)^{k+\kappa+s-1}|m|^{(k+2\kappa+s-2)}n^{(k+s-1)},\end{aligned}\tag{4.4.14}$$

As a shorthand let us set $F_{\mathrm{van}}=F_S^{1-\kappa(m)}(z,s)_{\mathrm{van}}$. Then we have the following expression for F_{van} as a triple sum

$$\begin{aligned}&(-1)^{1-\kappa}F_{\mathrm{van}}=\\&\sum_{n>0}\sum_{\llbracket\chi\rrbracket\in S}^{\lim}\sum_{k\geq 0}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\,\mathrm{e}(-n\llbracket\chi^{-1}\rrbracket\cdot\infty)(4\pi^2\varrho\llbracket\chi\rrbracket)^{k+\kappa+s-1}|m|^{(k+2\kappa+s-2)}n^{(k+s-1)}\,\mathrm{e}(nz)\end{aligned}\tag{4.4.15}$$

where $\sum_{\llbracket\chi\rrbracket\in S}^{\lim}$ is a shorthand for $\lim_{K\to\infty}\sum_{\llbracket\chi\rrbracket\in S_{\leq K}}$. We now move the summation over n past the other two summations, simultaneously pulling the terms $\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)$ outside the summation over k, and combining the exponents involving n. We thus obtain

$$\begin{aligned}&(-1)^{1-\kappa}F_{\mathrm{van}}=\\&\sum_{\llbracket\chi\rrbracket\in S}^{\lim}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\sum_{k\geq 0}(4\pi^2\varrho\llbracket\chi\rrbracket)^{k+\kappa+s-1}|m|^{(k+2\kappa+s-2)}\sum_{n>0}n^{(k+s-1)}\,\mathrm{e}(n(z-\llbracket\chi^{-1}\rrbracket\cdot\infty)).\end{aligned}\tag{4.4.16}$$

Applying the Lipschitz summation formula (3.3.1) to each summation over n in (4.4.16) we find that

$$\begin{aligned}&(-1)^{1-\kappa}F_{\mathrm{van}}=\\&\sum_{\llbracket\chi\rrbracket\in S}^{\lim}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)\sum_{k\geq 0}(4\pi^2\varrho\llbracket\chi\rrbracket)^{\kappa-1}|m|^{(k+2\kappa+s-2)}\sum_{n\in\mathbb{Z}}(2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket)^{k+s}(z+n-\chi^{-1}\rrbracket\cdot\infty)^{-k-s}\end{aligned}\tag{4.4.17}$$

where $\chi^{-1}\rrbracket$ is any (left) coset of $B(\mathbb{Z})$ in the double coset $\llbracket\chi^{-1}\rrbracket$, for each $\llbracket\chi\rrbracket\in S$. Using the identity (2.4.5) we write F_{van} as $F_{\mathrm{van}}=\sum_{\llbracket\chi\rrbracket\in S}^{\lim}\mathrm{e}(m\llbracket\chi\rrbracket\cdot\infty)F_{\mathrm{van},\llbracket\chi\rrbracket}$ where

$$F_{\mathrm{van},\llbracket\chi\rrbracket}=\sum_{k\geq 0}\sum_{n\in\mathbb{Z}}(-2\pi\mathbf{i}\varrho\llbracket\chi\rrbracket\,m)^{k+2\kappa+s-2}(z+n-\chi^{-1}\rrbracket\cdot\infty)^{-k-2\kappa-s+2}\,\mathrm{jac}(\llbracket\chi\,,z+n)^{1-\kappa}.\tag{4.4.18}$$

Recalling the definition of the generalized exponential function $\mathrm{e}(z,s)$ from (2.1.6) we see that

$$F_{\mathrm{van}} = \sum_{[\![\chi \in U^{\times}} \mathrm{e}(m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}\left(\frac{-\varrho\,[\![\chi]\!]\,m}{z-\chi^{-1}]\!]\cdot\infty}, s-2(1-\kappa)\right)\mathrm{jac}([\![\chi\ , z)^{1-\kappa}. \tag{4.4.19}$$

According to (2.4.4) we have $\varrho\,[\![\chi]\!]\,/(z-\chi^{-1}\cdot\infty) = \chi\cdot\infty - \chi\cdot z$. Using this identity together with the Kummer transformation $\Phi(a,b,z) = \mathrm{e}(z)\Phi(b-a,b,-z)$ (cf. (2.1.3)) we readily find that

$$\mathrm{e}(m\,[\![\chi]\!]\cdot\infty)\,\mathrm{e}\left(\frac{-\varrho\,[\![\chi]\!]\,m}{z-\chi^{-1}]\!]\cdot\infty}, s-2(1-\kappa)\right) = \mathrm{e}(m\,[\![\chi]\!]\cdot z)\,\mathrm{t}^{1-\kappa}(m,[\![\chi\ ,z,s), \tag{4.4.20}$$

so that we have

$$F_{\mathrm{van}} = \sum_{[\![\chi\in U^{\times}} T^{1-\kappa(m)}_{[\![\chi}(z,s). \tag{4.4.21}$$

This establishes the required identity (4.4.13). □

In the course of proving Theorem 4.4.5 we have also established the Fourier expansion of the continued Rademacher sums $T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$. We record the result as follows.

Theorem 4.4.6. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0$. Then we have*

$$T^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + F^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)_{\mathrm{van}} \tag{4.4.22}$$

for $z\in\mathbb{H}$ and $\Re(s)>1$.

Theorem 4.4.5 will facilitate the identification of the Fourier expansion of the modified Rademacher sum $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ obtained by sending s to 1 in $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z,s)$. Indeed, we will see presently (cf. Proposition 4.4.8) that the series $\tilde{F}^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(q)_{\mathrm{reg}}$ encodes the regular part of the fourier series expansion of $Q^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ precisely, in the case that $\kappa\leq 0$. In preparation for this we state the following result which identifies the value of the zeta function $Z^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(s)$ at $s=0$.

Proposition 4.4.7. *Let $S\subset[\![G(\mathbb{Q})]\!]^{\times}$, let $\kappa,m\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0$. Then we have*

$$Z^{\kappa(m)}_{S}(0) = \mathrm{c}^{\kappa}_{S}(m,0). \tag{4.4.23}$$

Proof. According to the proof of Lemma 4.3.1 the function $\Gamma(s)Z^{\kappa(m)}_{[\![\chi]\!]}(s)$ coincides with

$$(-1)^{1-\kappa}\,\mathrm{e}(m\,[\![\chi]\!]\cdot\infty)(4\pi^2\varrho\,[\![\chi]\!])^{1-s-\kappa}|m|^{(1-s-2\kappa)}(F(\alpha_{[\![\chi]\!]},s)+\mathrm{e}(-s/2)F(-\alpha_{[\![\chi]\!]},s)) \tag{4.4.24}$$

where $0 \leq \alpha_{[\![\chi]\!]} < 1$ and $\alpha_{[\![\chi]\!]} + \mathbb{Z} = -\chi^{-1} \cdot \infty + \mathbb{Z}$. We have $F(\alpha, 0) + F(-\alpha, 0) = -1$ for $\alpha \in \mathbb{R}$ (cf. [KR01]), so taking $s = 0$ in (4.4.24) we obtain

$$Z_{[\![\chi]\!]}^{\kappa(m)}(0) = \mathrm{e}(m\,[\![\chi]\!] \cdot \infty)(-4\pi^2 \varrho\,[\![\chi]\!])^{1-\kappa} m^{(1-2\kappa)}. \tag{4.4.25}$$

Upon inspection of (3.2.21) we see that

$$\mathrm{c}_S^{\kappa}(m, 0) = \sum_{[\![\chi]\!] \in S} \mathrm{e}(m\,[\![\chi]\!] \cdot \infty)(-4\pi^2 \varrho\,[\![\chi]\!])^{1-\kappa} m^{(1-2\kappa)} \tag{4.4.26}$$

for $\kappa \leq 0$, and the claim now follows since $Z_S^{\kappa(m)}(s) = \sum_{[\![\chi]\!] \in S} Z_{[\![\chi]\!]}^{\kappa(m)}(s)$ by definition. □

Proposition 4.4.8. *Let Γ be a group commensurable with $G(\mathbb{Z})$, and let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then we have*

$$Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)_{\mathrm{reg}} \tag{4.4.27}$$

for $z \in \mathbb{H}$.

Proof. We take the limit as s tends to 1 in the right hand side of the identity (4.4.11). We have $F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z, 1)_{\mathrm{van}} = F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)_{\mathrm{van}}$ by definition (cf. §4.2), and we have $Z_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(0) = \mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa}(m, 0)$ by Proposition 4.4.7. The required identity (4.4.27) now follows from the fact that $F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)_{\mathrm{reg}} = \mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa}(m, 0) + F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)_{\mathrm{van}}$, also by definition (cf. §3.2). □

The proof of Proposition 4.4.8 implies a result analogous to (4.4.27) for the normalized Rademacher sum $T_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$.

Proposition 4.4.9. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then we have*

$$T_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(mz) + F_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)_{\mathrm{van}} \tag{4.4.28}$$

for $z \in \mathbb{H}$.

We may compare the functions $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ and $R_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ defined by the modified and classical Rademacher sums, respectively. Inspecting Theorem 3.3.2 and Proposition 4.4.8 we obtain the precise relationship, which we record in the following proposition.

Proposition 4.4.10. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then we have*

$$Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z) = R_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z) + \frac{1}{2}\mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa}(m, 0). \tag{4.4.29}$$

According to Propositions 4.4.8 and 4.4.9 the modified Rademacher sum $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ and the normalized Rademacher sum $T_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ also differ only by a constant function.

Proposition 4.4.11. *Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ be cusps of Γ. Let $\kappa,m\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0$. Then we have*

$$Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)=T_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)+\mathrm{c}_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa}(m,0). \tag{4.4.30}$$

4.5 Variance

In this section we examine how the modified Rademacher sum $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ transforms under the weight 2κ action of Γ. Combining Theorem 3.4.4 and Proposition 4.4.10 we obtain the result that the modified Rademacher sum $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ is an automorphic integral of weight 2κ for Γ in case Γ has width one at infinity.

Theorem 4.5.1. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p}\in\mathsf{P}_\Gamma$ be a cusp of Γ. Let $\kappa,m\in\mathbb{Z}$ such that $\kappa\leq 0$ and $m<0$. Then the modified Rademacher sum $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ is an automorphic integral of weight 2κ for Γ, and for $\mathsf{q}\in\mathsf{P}_\Gamma$ another cusp of Γ the Fourier expansion of the function $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z)$ is the expansion of $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ at the cusp q.*

The proof of Theorem 3.4.4 yields the following explicit description of the associated cusp form map $I_\kappa(\Gamma)\to S_{1-\kappa}(\Gamma)$ on the subspace of $I_\kappa(\Gamma)$ spanned by the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$.

Proposition 4.5.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p}\in\mathsf{P}_\Gamma$ be a cusp of Γ. Let $\kappa,m\in\mathbb{Z}$ such that $\kappa\leq 0<m$. Then*

$$m^{\kappa}Q_{\Gamma,\mathsf{p}}^{\kappa(-m)}\mapsto m^{1-\kappa}P_{\Gamma,\mathsf{p}}^{1-\kappa(m)}\mapsto 0, \tag{4.5.1}$$

under the maps $I_\kappa(\Gamma)\mapsto S_{1-\kappa}(\Gamma)$ of §2.7.

We next seek to establish the utility of Proposition 4.5.2 by identifying the subspace of $I_\kappa(\Gamma)$ that is spanned by the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$.

Theorem 4.5.3. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity, and let $\kappa\in\mathbb{Z}$ with $\kappa\leq 0$. Then the set $\{Q_{\Gamma,\mathsf{p}}^{\kappa(-m)}(z)\mid \mathsf{p}\in\mathsf{P}_\Gamma, m\in\mathbb{Z}^+\}$ is a basis for the space of automorphic integrals of weight 2κ for Γ in case $\kappa<0$. When $\kappa=0$ the set $\{Q_{\Gamma,\mathsf{p}}^{(-m)}(z)\mid \mathsf{p}\in\mathsf{P}_\Gamma, m\in\mathbb{Z}^+\}$ spans a subspace of $I_0(\Gamma)$ of codimension 1.*

Proof. For $m>0$ the function $Q_{\Gamma,\mathsf{p}}^{\kappa(-m)}(z)$ has principal part $\mathrm{e}(-mz)$ at the cusp p, and vanishes at all the other cusps, so the collection $\{Q_{\Gamma,\mathsf{p}}^{\kappa(-m)}(z)\mid \mathsf{p}\in\mathsf{P}_\Gamma, m\in\mathbb{Z}^+\}$ is linearly independent. To show that it furnishes a basis we follow Niebur's proof of Theorem 3.3 in [Nie74], employing

Petersson's generalized Riemann–Roch Theorem to show that the dimension of the subspace of $I_\kappa(\Gamma)$ containing automorphic integrals with poles of degree not more than m say at p is bounded above by m less the dimension of the space $S_{1-\kappa}(\Gamma)$, for m sufficiently large. □

We write $I_0'(\Gamma)$ for the subspace of $I_0(\Gamma)$ spanned by the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{(-m)}(z)$ of weight 0 for varying $\mathsf{p} \in \mathsf{P}_\Gamma$ and $m \in \mathbb{Z}^+$.

Theorem 4.5.4. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity. Then for any $\kappa \in \mathbb{Z}$ the sequence*

$$0 \to M_\kappa(\Gamma) \to I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma) \to 0 \tag{4.5.2}$$

is exact.

Proof. The second map of (4.5.2) is an inclusion by definition, and we observed already in §2.7 that $M_\kappa(\Gamma)$ is the kernel of the associated cusp form map $I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma)$ (cf. Lemma 2.7.1), so we require to show only that the second to last map of (4.5.2) is surjective. For this we recall the fact (cf. [Iwa97]) that the holomorphic Poincaré series $P_{\Gamma,\mathsf{p}}^{1-\kappa(m)}(z)$, with varying $m \in \mathbb{Z}^+$, span the space $S_{1-\kappa}(\Gamma)$ of cusp forms of weight $2-2\kappa$ for Γ whenever $\kappa \leq 0$. Indeed, one need not vary the cusp p in order to obtain a spanning set. Given an arbitrary cusp form $g \in S_{1-\kappa}(\Gamma)$ we may then write

$$g = \sum_{n>0} a_n P_\Gamma^{1-\kappa(n)} \tag{4.5.3}$$

for some $a_n \in \mathbb{C}$, with only finitely many a_n non-zero. Employing Proposition 4.5.2 we see that the automorphic integral

$$f = \sum_{n>0} n^{2\kappa-1} a_n Q_\Gamma^{\kappa(-n)} \tag{4.5.4}$$

is mapped to g by the associated cusp form map. We conclude that the associated cusp form map is surjective. This completes the proof. □

Theorem 4.5.3 and Proposition 4.5.2 together describe the associated cusp form map $I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma)$ explicitly, and thus we obtain a powerful criterion for determining when an automorphic integral $f \in I_\kappa(\Gamma)$ lies in the subspace $M_\kappa(\Gamma)$ of modular forms. For example, we have the following corollary, indicating exactly when the modified Rademacher sum $Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z)$ is a modular form for Γ in the sense of §2.7.

Corollary 4.5.5. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the modified*

Rademacher sum $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ *is a modular form of weight* 2κ *for* Γ *if and only if the cusp form* $P^{1-\kappa(-m)}_{\Gamma,\mathsf{p}}(z)$ *vanishes identically.*

5 Structural applications

We have seen already in §4 various applications of the modified Rademacher sums, and the continuation procedure introduced in §4.1, such as the basis theorem (Theorem 4.5.3) for automorphic integrals, and the explicit description (cf. Proposition 4.5.2) of the associated cusp form map $I_\kappa(\Gamma) \to S_{1-\kappa}(\Gamma)$. In this section we explore further consequences of the modified Rademacher sum construction for the structure of the spaces of automorphic integrals associated to groups commensurable with the modular group. Applications of the modified Rademacher sums to monstrous moonshine and quantum gravity will be developed in §§6,7.

5.1 Constants

Perhaps the most striking application of the continuation procedure of §4.1 is the correction of the constant terms appearing in the classical Rademacher sums. For negative weights, the constant term in the Fourier expansion of an automorphic integral $f \in I_\kappa(\Gamma)$ is determined by the automorphy condition (2.7.13). At weight 0 constant functions are themselves automorphic integrals, so the classical Rademacher sums $R^{(m)}_{\Gamma,\mathsf{p}}(z)$ of weight 0 do not fail to be automorphic, and it is natural then to wonder about the significance of the constant term $\mathrm{c}_{\Gamma,\mathsf{p}}(m,0)$ appearing in the Fourier expansion of the modified Rademacher sum $Q^{(m)}_{\Gamma,\mathsf{p}}(z)$ of weight 0 associated to a group Γ at a cusp $\mathsf{p} \in \mathsf{P}_\Gamma$.

Let $\mathfrak{G}$ denote the set of subgroups of $G(\mathbb{Q})$ that are commensurable with $G(\mathbb{Z})$. Recall from §2.7 that for $\Gamma \in \mathfrak{G}$ we write $M_0(\Gamma)$ for the space of holomorphic functions on $\mathbb{H}$ that are invariant for the natural action of Γ and are meromorphic at the cusps of Γ. Let $\mathfrak{M}_0$ be the union of the spaces $M_0(\Gamma)$ for $\Gamma \in \mathfrak{G}$.

$$\mathfrak{M}_0 = \bigcup_{\Gamma \in \mathfrak{G}} M_0(\Gamma) \tag{5.1.1}$$

Since $\mathfrak{G}$ is closed under intersections, the set $\mathfrak{M}_0$ is in fact a subalgebra of the space $\mathcal{O}(\mathbb{H})$ of holomorphic functions on $\mathbb{H}$. Since $G(\mathbb{Q})$ is the commensurator of $G(\mathbb{Z})$, the algebra $\mathfrak{M}_0$ is $G(\mathbb{Q})$-invariant. The constant functions evidently constitute a $G(\mathbb{Q})$-invariant subspace $\mathbb{C}1 \subset \mathfrak{M}_0$. We may enquire as to the $G(\mathbb{Q})$-module structure of the quotient $\mathfrak{M}_0/\mathbb{C}1$. According to [Nor84] this quotient is irreducible, and the natural map $\mathfrak{M}_0 \to \mathfrak{M}_0/\mathbb{C}1$ admits a section.

Theorem 5.1.1 ([Nor84]). *There exists a unique* $G(\mathbb{Q})$*-submodule* $\mathfrak{M}_0' \subset \mathfrak{M}_0$ *with the property* *that the natural inclusions* $\mathbb{C}1, \mathfrak{M}_0' \subset \mathfrak{M}_0$ *induce an isomorphism*

$$\mathfrak{M}_0 \cong \mathbb{C}1 \oplus \mathfrak{M}_0' \tag{5.1.2}$$

of $G(\mathbb{Q})$*-modules.*

According to Theorem 5.1.1, for each $f \in \mathfrak{M}_0$ there is a unique $c(f) \in \mathbb{C}$ with the property that $f + c(f) \in \mathfrak{M}_0'$. The value $c(f)$ is called the *Rademacher constant of* f. The assignment $f \mapsto c(f)$ defines a linear function on $\mathfrak{M}_0$ with kernel $\mathfrak{M}_0'$ which we call the *Rademacher constant function*. We suggest that the Rademacher constant function may be generalized to automorphic integrals of weight 0 in the following way. Let $\mathfrak{I}_0$ be the union of the spaces $I_0(\Gamma)$ for $\Gamma \in \mathfrak{G}$.

$$\mathfrak{I}_0 = \bigcup_{\Gamma \in \mathfrak{G}} I_0(\Gamma) \tag{5.1.3}$$

Since $\mathfrak{G}$ is closed under intersections, the set $\mathfrak{I}_0$ is also a subalgebra of $\mathcal{O}(\mathbb{H})$, and since $G(\mathbb{Q})$ is the commensurator of $G(\mathbb{Z})$, the algebra $\mathfrak{I}_0$ is also stable under the action of $G(\mathbb{Q})$. The constant functions again constitute a $G(\mathbb{Q})$-invariant subspace $\mathbb{C}1 \subset \mathfrak{I}_0$. We conjecture that the natural analogue of Theorem 5.1.1 holds for $\mathfrak{I}_0$. Precisely, we conjecture that there is a unique $G(\mathbb{Q})$-submodule $\mathfrak{I}_0' \subset \mathfrak{I}_0$ with the property that the natural inclusions induce an isomorphism

$$\mathfrak{I}_0 \cong \mathbb{C}1 \oplus \mathfrak{I}_0' \tag{5.1.4}$$

of $G(\mathbb{Q})$-modules. Further, we conjecture that $\mathfrak{I}_0'$ is just the space spanned by the modified Rademacher sums $Q_{\Gamma,\mathsf{p}}^{\kappa(-m)}(z)$ for $\Gamma \in \mathfrak{G}$ and $\mathsf{p} \in \mathsf{P}_\Gamma$ and $m \in \mathbb{Z}^+$.

5.2 Inner products

Let Γ be a group commensurable with $G(\mathbb{Z})$ and let $\kappa \in \mathbb{Z}^+$. Then the space $S_\kappa(\Gamma)$ of cusp forms of weight 2κ for Γ becomes a Hilbert space when equipped with the *Petersson inner product*, defined by setting

$$\langle f, g \rangle = \int_{\mathfrak{F}_\Gamma} f(z)\overline{g(z)}\Im(z)^{2\kappa} \mathrm{d}\mu(z) \tag{5.2.1}$$

for $f, g \in S_\kappa(\Gamma)$, where $\mathfrak{F}_\Gamma$ is a fundamental domain for Γ. Let us define a *normalized Petersson inner product of weight* 2κ *for* Γ, to be denoted $\langle \cdot\,, \cdot \rangle_\Gamma^\kappa$, by setting

$$\langle f, g \rangle_\Gamma^\kappa = (4\pi)^{(2\kappa-2)} \int_{\mathfrak{F}_\Gamma} f(z)\overline{g(z)}\Im(z)^{2\kappa} \mathrm{d}\mu(z) \tag{5.2.2}$$

for $f,g\in S_\kappa(\Gamma)$. It is well known (and follows from Theorem 3.4.3), that for $\kappa,m\in\mathbb{Z}^+$ and p a cusp for Γ, the holomorphic Poincaré series $P^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ is a cusp form of weight 2κ for Γ in case Γ has width one at infinity. A standard calculation (cf. [Iwa97]) shows that we have

$$\left\langle P^{\kappa(m)}_{\Gamma,\mathsf{p}},P^{\kappa(n)}_{\Gamma,\mathsf{q}}\right\rangle^{\kappa}_{\Gamma}=n^{2-2\kappa}(\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\delta_{m,n}+\mathrm{c}^{\kappa}_{\Gamma,\mathsf{p}|\mathsf{q}}(m,n))\tag{5.2.3}$$

for $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ and $m,n\in\mathbb{Z}^+$. According to Theorem 3.4.3 the right hand side of (5.2.3) is, up to the scalar factor $n^{2-2\kappa}$, the coefficient of $\mathrm{e}(nz)$ in the expansion $P^{\kappa(m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ of the Poincaré series $P^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ at q. Now the modified Rademacher sums $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$, for $-\kappa,-m\in\mathbb{Z}^+$, span the space $I_\kappa(\Gamma)$ of automorphic integrals for weight 2κ for Γ, according to Theorem 4.5.3, and thus serve an analogous rôle for the spaces $I_\kappa(\Gamma)$ as the Poincaré series $P^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ do for the spaces $S_\kappa(\Gamma)$. It is natural then to use the right hand side of (5.2.3) to extend the normalized Petersson inner product to negative weights. In light of the fact, established in Proposition 4.4.8, that the coefficient of $\mathrm{e}(-nz)$ in the expansion $Q^{\kappa(-m)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ of the modified Rademacher sum $Q^{\kappa(-m)}_{\Gamma,\mathsf{p}}(z)$ at q is $\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\delta_{m,n}$ for $m,n\in\mathbb{Z}^+$, we define an inner product $\langle\cdot\,,\cdot\rangle^{\kappa}_{\Gamma}$ on $I_\kappa(\Gamma)$ by setting

$$\left\langle Q^{\kappa(-m)}_{\Gamma,\mathsf{p}},Q^{\kappa(-n)}_{\Gamma,\mathsf{q}}\right\rangle^{\kappa}_{\Gamma}=n^{2-2\kappa}\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\delta_{m,n}\tag{5.2.4}$$

for $\mathsf{p},\mathsf{q}\in\mathsf{P}_\Gamma$ and $m,n\in\mathbb{Z}^+$, when $\kappa<0$. In case $\kappa=0$ the modified Rademacher sums $Q^{\kappa(-m)}_{\Gamma,\mathsf{p}}(z)$ span a subspace $I_0'(\Gamma)$ of $I_0(\Gamma)$ of codimension 1 (cf. Theorem 4.5.3 and §5.1) and we may use the formula (5.2.4) to define an inner product on this subspace $I_0'(\Gamma)$.

$$\left\langle Q^{(-m)}_{\Gamma,\mathsf{p}},Q^{(-n)}_{\Gamma,\mathsf{q}}\right\rangle^{0}_{\Gamma}=n^{2}\delta_{\Gamma,\mathsf{p}|\mathsf{q}}\delta_{m,n}\tag{5.2.5}$$

A complement to $I_0'(\Gamma)$ in $I_0(\Gamma)$ is spanned by the constant functions. We extend the inner product $\langle\cdot\,,\cdot\rangle^{0}_{\Gamma}$ to all of $I_0(\Gamma)$ by adopting the convention that

$$\left\langle Q^{(-m)}_{\Gamma,\mathsf{p}},f\right\rangle^{0}_{\Gamma}=0\tag{5.2.6}$$

for all $\mathsf{p}\in\mathsf{P}_\Gamma$ and $m\in\mathbb{Z}^+$ whenever f is identically constant.

5.3 Branching

Let Δ and Γ be groups commensurable with $G(\mathbb{Z})$ and suppose that Δ is a subgroup of Γ. Then an automorphic integral for Γ is also an automorphic integral for Δ. By Theorem 4.5.3 the modified Rademacher sums span the spaces of automorphic integrals, so it is natural to consider the problem of expressing the modified Rademacher sums $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}$ associated to Γ in terms of the modified Rademacher sums $Q^{\kappa(n)}_{\Delta,\mathsf{r}}$ associated to Δ.

Suppose then that Δ and Γ are groups commensurable with $G(\mathbb{Z})$ and Δ is a subgroup of

Γ. Let p be a cusp for Γ, let Σ_{p} be a scaling coset for Γ at p, and choose a system $\{\Sigma_{\mathsf{r}} \mid \mathsf{r} \in \mathsf{P}_\Delta\}$ of scaling cosets for Δ. Suppose for now that κ is a negative integer. Then by Theorem 4.5.3 we have

$$Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z) = \sum_{\mathsf{r}\in\mathsf{P}_\Delta} b^n_{\mathsf{r}} Q^{\kappa(n)}_{\Delta,\mathsf{r}}(z) \tag{5.3.1}$$

for some $b^n_{\mathsf{r}} \in \mathbb{C}$, and these b^n_{r} are non-zero for only finitely many $n < 0$. Furthermore, since the modified Rademacher sum $Q^{\kappa(m)}_{\Gamma,\mathsf{p}}(z)$ has no poles away from p, the coefficient b^n_{r} can be non-zero only when r lies in the preimage of p under the natural map $\mathsf{X}_\Delta \to \mathsf{X}_\Gamma$, which is to say, b^n_{r} is non-zero only when $\mathsf{r} \in \Delta\backslash\mathsf{p}$. It is natural then to consider the map

$$\begin{aligned} \Delta\backslash\Sigma_{\mathsf{p}} &\to \Delta\backslash\mathsf{p} \\ \Delta\sigma &\mapsto \Delta\sigma\cdot\infty \end{aligned} \tag{5.3.2}$$

sending right cosets of Δ in Σ_{p} to cusps of Δ contained in $\mathsf{p} \subset \hat{\mathbb{Q}}$. The map (5.3.2) is always surjective. We may ask under what circumstances it is also injective.

Lemma 5.3.1. *Let Δ and Γ be groups commensurable with $G(\mathbb{Z})$ and suppose that Δ is a subgroup of Γ. Let p be a cusp for Γ and let Σ_{p} be a scaling coset for Γ at p. Then the map $\Delta\backslash\Sigma_{\mathsf{p}} \to \Delta\backslash\mathsf{p}$ of (5.3.2) is injective if and only if $\Delta\sigma$ is a scaling coset for Δ at $\Delta\sigma\cdot\infty$ for every $\Delta\sigma \in \Delta\backslash\Sigma_{\mathsf{p}}$.*

Proof. Suppose the map $\Delta\backslash\Sigma_{\mathsf{p}} \to \Delta\backslash\mathsf{p}$ of (5.3.2) is injective, and let $\sigma \in \Sigma_{\mathsf{p}}$. The forward implication of the lemma follows if we can show that $\Delta\sigma$ is a scaling coset for Δ at $\Delta\sigma\cdot\infty$. For this it suffices to show that $(\sigma^{-1}\Delta\sigma)_\infty = B(\mathbb{Z})$. Since $\sigma^{-1}\Delta\sigma$ is a subgroup of $\sigma^{-1}\Gamma\sigma$, and $(\sigma^{-1}\Gamma\sigma)_\infty = (\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}})_\infty = B(\mathbb{Z})$ by the defining properties of Σ_{p}, we have the inclusion $(\sigma^{-1}\Delta\sigma)_\infty \subset B(\mathbb{Z})$, so it suffices to show that the translation T belongs to $\sigma^{-1}\Delta\sigma$. Now $\Delta\sigma T$ is also a coset of Δ in Σ_{p}, since $\Gamma\sigma = \Sigma_{\mathsf{p}}$ is a union of left cosets of $B(\mathbb{Z})$. Since T fixes ∞ we have $\Delta\sigma T\cdot\infty = \Delta\sigma\cdot\infty$, and thus the cosets $\Delta\sigma T$ and $\Delta\sigma$ coincide by the assumed injectivity of the map (5.3.2). It follows that $T \in \sigma^{-1}\Delta\sigma$, so that $\Delta\sigma$ is indeed a scaling coset for Δ.

For the reverse implication suppose that $\Delta\sigma$ is a scaling coset for Δ for every $\sigma \in \Sigma_{\mathsf{p}}$. Let $\sigma, \sigma' \in \Sigma_{\mathsf{p}}$ and suppose that $\Delta\sigma\cdot\infty = \Delta\sigma'\cdot\infty$. Then $\sigma^{-1}\delta\sigma' \in (\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}})_\infty = B(\mathbb{Z})$ for some $\delta \in \Delta$, so $\sigma^{-1}\delta\sigma' = T^n$ for some $n \in \mathbb{Z}$, and this implies $\Delta\sigma' = \Delta\sigma T^n$. Now the assumption that $\Delta\sigma$ is a scaling coset for Δ implies that $\Delta\sigma$ is a union of left cosets of $B(\mathbb{Z})$, so $\Delta\sigma T^n = \Delta\sigma$. We conclude that $\Delta\sigma' = \Delta\sigma$, so that the map $\Delta\backslash\Sigma_{\mathsf{p}} \to \Delta\backslash\mathsf{p}$ is indeed injective. This completes the proof of the lemma. ☐

According to the discussion of §2.6 a coset $\Delta\sigma \in \Delta\backslash\Sigma_{\mathsf{p}}$ becomes a scaling coset for Δ at $\Delta\sigma\cdot\infty$ only once we multiply it on the right by $[\beta]$ (cf. (2.3.3)) for some $\beta \in \mathbb{Q}^+$. In fact, this β is a positive integer, for we have $(\sigma^{-1}\Delta\sigma)_\infty \subset (\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}})_\infty = B(\mathbb{Z})$, so that $(\sigma^{-1}\Delta\sigma)_\infty$ is generated by T^n for some $n \in \mathbb{Z}^+$. In order that $\Delta\sigma[\beta]$ be a scaling coset for Δ at $\Delta\sigma \cdot \infty$ we should have $([1/\beta]\sigma^{-1}\Delta\sigma[\beta])_\infty = B(\mathbb{Z})$, so the computation $[1/\beta]T^n[\beta] = T^{n/\beta}$ shows that $\beta = n$. Evidently we may define a function $\Delta\backslash\Sigma_{\mathsf{p}} \to \mathbb{Z}^+$ by mapping the coset $\Delta\sigma$ to the positive integer n such that T^n generates $(\sigma^{-1}\Delta\sigma)_\infty$. We next show that this function $\Delta\backslash\Sigma_{\mathsf{p}} \to \mathbb{Z}^+$ factors through the map $\Delta\backslash\Sigma_{\mathsf{p}} \to \Delta\backslash\mathsf{p}$ of (5.3.2).

Lemma 5.3.2. *Let Δ and Γ be groups commensurable with $G(\mathbb{Z})$, and suppose that Δ is a subgroup of Γ. Let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ and let Σ_{p} be a scaling coset for Γ at p. Then for any $\sigma_1, \sigma_2 \in \Sigma_{\mathsf{p}}$ the identity $\Delta\sigma_1 \cdot \infty = \Delta\sigma_2 \cdot \infty$ implies $(\sigma_1^{-1}\Delta\sigma_1)_\infty = (\sigma_2^{-1}\Delta\sigma_2)_\infty$.*

Proof. Set $\mathsf{r} = \Delta\sigma_1 \cdot \infty = \Delta\sigma_2 \cdot \infty$. Since $(\sigma^{-1}\Gamma\sigma)_\infty = (\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}})_\infty = B(\mathbb{Z})$ for any $\sigma \in \Sigma_{\mathsf{p}}$ the groups $(\sigma_1^{-1}\Delta\sigma_1)_\infty$ and $(\sigma_2^{-1}\Delta\sigma_2)_\infty$ are both contained in $B(\mathbb{Z})$. Let n_1 and n_2 be the positive integers such that $(\sigma_1^{-1}\Delta\sigma_1)_\infty$ is generated by T^{n_1} and $(\sigma_2^{-1}\Delta\sigma_2)_\infty$ is generated by T^{n_2}. Then $\Delta\sigma_1[n_1]$ and $\Delta\sigma_2[n_2]$ are both scaling cosets for Δ at r, so $\Delta\sigma_1[n_1] = \Delta\sigma_2[n_2]T^\alpha$ for some $\alpha \in \mathbb{Q}$. This implies that $[n_2]T^\alpha[1/n_1]$ is an element of $(\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}})_\infty = B(\mathbb{Z})$, so that $[n_2]T^\alpha[1/n_1] = T^k$ for some $k \in \mathbb{Z}$. Now we compute

$$[n_2]T^\alpha[1/n_1] = \begin{bmatrix} n_2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & n_1 \end{bmatrix} = \begin{bmatrix} n_2 & n_2\alpha n_1 \\ 0 & n_1 \end{bmatrix} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \tag{5.3.3}$$

and conclude that $n_1 = n_2$, as we required to show. □

On the basis of Lemma 5.3.2 we may define a function $\Delta\backslash\mathsf{p} \to \mathbb{Z}^+$ for each cusp p of Γ by first choosing a scaling coset Σ_{p} for Γ at p, and then sending the cusp $\mathsf{r} \in \Delta\backslash\mathsf{p}$ of Δ to the unique positive integer n satisfying $(\sigma^{-1}\Delta\sigma)_\infty = \langle T^n \rangle$ when $\mathsf{r} = \Delta\sigma \cdot \infty$. Observe now that this map does not depend upon the choice of scaling coset Σ_{p}, for if we replace $\sigma \in \Sigma_{\mathsf{p}}$ with $\sigma' = \sigma T^\alpha \in \Sigma_{\mathsf{p}}T^\alpha$ for some $\alpha \in \mathbb{Q}$, then we have

$$((\sigma')^{-1}\Delta\sigma')_\infty = (T^{-\alpha}\sigma^{-1}\Delta\sigma T^\alpha)_\infty = T^{-\alpha}(\sigma^{-1}\Delta\sigma)_\infty T^\alpha = (\sigma^{-1}\Delta\sigma)_\infty. \tag{5.3.4}$$

In light of this we may define a map $w_\Gamma : \mathsf{P}_\Delta \to \mathbb{Z}^+$, which we call the *width function associated to* Γ, by setting $w_\Gamma(\mathsf{r}) = n$ in case $\mathsf{r} = \Delta\sigma \cdot \infty$ and T^n generates $(\sigma^{-1}\Delta\sigma)_\infty$ and $\Gamma\sigma$ is a scaling coset for Γ at the unique cusp of Γ containing r.

$$\begin{gathered} w_\Gamma : \mathsf{P}_\Delta \to \mathbb{Z}^+ \\ \mathsf{r} = \Delta\sigma \cdot \infty,\, \Gamma\sigma \in \mathfrak{S}_{\Gamma\mathsf{r}} \Longrightarrow (\sigma^{-1}\Delta\sigma)_\infty = \langle T^{w_\Gamma(\mathsf{r})} \rangle \end{gathered} \tag{5.3.5}$$

We call $w_\Gamma(\mathsf{r})$ the *width of* r *with respect to* Γ. The next result gives upper bounds for the width functions.

Lemma 5.3.3. *Let* Δ *and* Γ *be groups commensurable with* $G(\mathbb{Z})$*, and suppose that* Δ *is a subgroup of* Γ*. Then we have*

$$\sum_{\mathsf{r}\in\Delta\backslash\mathsf{p}} w_\Gamma(\mathsf{r}) = \#\Delta\backslash\Gamma \tag{5.3.6}$$

for every cusp p *of* Γ*.*

Proof. Let p be a cusp of Γ and choose a scaling coset Σ_p for Γ at p, and let us temporarily write c_p for the map $\Delta\sigma \mapsto \Delta\sigma\cdot\infty$ of (5.3.2). We claim that $w_\Gamma(\mathsf{r})$ is the cardinality of the preimage of r under c_p, for each $\mathsf{r}\in\Delta\backslash\mathsf{p}$. The identity (5.3.6) follows from this because the cardinality $\#\Delta\backslash\Sigma_\mathsf{p}$ of the source of c_p is just the number of cosets of Δ in Γ. To compute $\#c_\mathsf{p}^{-1}(\mathsf{r})$ suppose $\Delta\sigma\cdot\infty = \Delta\sigma'\cdot\infty$ for some $\sigma,\sigma'\in\Sigma_\mathsf{p}$. Then $\sigma^{-1}\delta\sigma' \in (\Sigma_\mathsf{p}^{-1}\Sigma_\mathsf{p})_\infty = B(\mathbb{Z})$ for some $\delta\in\Delta$, so that $\Delta\sigma' = \Delta\sigma T^k$ for some $k\in\mathbb{Z}$. Now $\Delta\sigma T$ is an element of $\Delta\backslash\Sigma_\mathsf{p}$ whenever $\Delta\sigma$ is, for Σ_p is a union of left cosets of $B(\mathbb{Z})$ by construction. So we have $c_\mathsf{p}^{-1}(\mathsf{r}) = \{\Delta\sigma T^k \mid k\in\mathbb{Z}\}$ for any $\sigma\in\Sigma_\mathsf{p}$ satisfying $\mathsf{r} = \Delta\sigma\cdot\infty$. If $w_\Gamma(\mathsf{r}) = n$, so that $(\sigma^{-1}\Delta\sigma)_\infty = \langle T^n\rangle$ then $\Delta\sigma T^k = \Delta\sigma T^l$ if and only if $n|(k-l)$. This establishes the claim that $\#c_\mathsf{p}^{-1}(\mathsf{r}) = w_\Gamma(\mathsf{r})$, and completes the proof. □

We can now compute the coefficients b_r^n in (5.3.1) explicitly, and thus obtain a branching theorem for the modified Rademacher sums. Even though we formulated the expression (5.3.1) under the assumption that κ be negative, our methods will apply for all $\kappa\in\mathbb{Z}$ such that $\kappa\le 0$.

Theorem 5.3.4. *Let* Δ *and* Γ *be groups commensurable with* $G(\mathbb{Z})$*. Suppose that* Δ *is a subgroup of* Γ*, and suppose that* Δ *and* Γ *both have width one at infinity. Let* p *be a cusp of* Γ *and let* Σ_p *be a scaling coset for* Γ *at* p*. Let* $\kappa, m\in\mathbb{Z}$ *such that* $\kappa\le 0$ *and* $m<0$*. Then there exists a system of scaling cosets* $\{\Sigma_\mathsf{r}\mid \mathsf{r}\in\mathsf{P}_\Delta\}$ *for* Δ *for which we have*

$$Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z) = \sum_{\mathsf{r}\in\Delta\backslash\mathsf{p}} w_\Gamma(\mathsf{r}) Q_{\Delta,\mathsf{r}}^{\kappa(mw_\Gamma(\mathsf{r}))}(z). \tag{5.3.7}$$

Proof. Let Δ and Γ be as in the statement of the lemma. Under the assumption that Γ has width one at infinity, and the convention that $\Sigma_{\Gamma\cdot\infty} = \Gamma$ in this case (cf. §2.6), we have

$$Q_{\Gamma,\mathsf{p}}^{\kappa(m)}(z) = Q_{[\![\Sigma_\mathsf{p}^{-1}}^{\kappa(m)}(z) = \sum_{\Delta\sigma\in\Delta\backslash\Sigma_\mathsf{p}} Q_{[\![\sigma^{-1}\Delta}^{\kappa(m)}(z), \tag{5.3.8}$$

and we may attempt to write each summand $Q_{[\![\sigma^{-1}\Delta}^{\kappa(m)}(z)$ in terms of the modified Rademacher sums $Q_{\Delta,\mathsf{r}}^{\kappa(n)}(z)$ associated to Δ. Let $\Delta\sigma\in\Delta\backslash\Sigma_\mathsf{p}$ and set $\mathsf{r} = \Delta\sigma\cdot\infty$. Supposing that $w_\Gamma(\mathsf{r}) = n$, so that $(\sigma^{-1}\Delta\sigma)_\infty = \langle T^n\rangle$, we set $\Sigma_{\mathsf{r}} = \Delta\sigma[n]$. Then we have

$$Q^{\kappa(m)}_{[\![\sigma^{-1}\Delta}(z) = Q^{\kappa(m)}_{[\![[n][1/n]\sigma^{-1}\Delta}(z) = Q^{\kappa(mn)}_{[\![\Sigma_{\mathsf{r}}^{-1}}(z), \tag{5.3.9}$$

since $Q^{\kappa(m)}_{[\![[n]\chi}(z,s) = Q^{\kappa(mn)}_{[\![\chi}(z,s)$ (cf. §4.1). Under the assumption that Δ has width one at infinity, so that $\Sigma_{\Delta\cdot\infty} = \Delta$, we may rewrite the right most term in (5.3.9) as $Q^{\kappa(mn)}_{\Delta,\mathsf{r}}(z)$. We see then that a coset $\Delta\sigma$ of Δ in Σ_{p} contributes a term $Q^{\kappa(mw_\Gamma(\mathsf{r}))}_{\Delta,\mathsf{r}}(z)$ to the right hand side of (5.3.7), where $\mathsf{r} = \Delta\sigma\cdot\infty$. According to the proof of Lemma 5.3.3 there are exactly $w_\Gamma(\mathsf{r})$ cosets $\Delta\sigma$ in $\Delta\backslash\Sigma_{\mathsf{p}}$ satisfying $\mathsf{r} = \Delta\sigma\cdot\infty$. The required identity (5.3.7) now follows. □

From the proof of Theorem 5.3.4 we see that the existence of the scaling cosets of the conclusion is verified constructively: if $\mathsf{r} = \Delta\sigma\cdot\infty$ for some coset $\Delta\sigma \in \Delta\backslash\Sigma_{\mathsf{p}}$, then we may take $\Sigma_{\mathsf{r}} = \Delta\sigma[w_\Gamma(\mathsf{r})]$.

By utilizing the inner products $\langle\cdot\,,\cdot\rangle^\kappa_\Delta$ of §5.2 we can reformulate Theorem 5.3.4 in a way that is independent of scaling coset choices.

Theorem 5.3.5. *Let Δ and Γ be groups commensurable with $G(\mathbb{Z})$. Suppose that Δ is a subgroup of Γ, and suppose that Δ and Γ both have width one at infinity. Let p be a cusp of Γ, let κ min$\mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then we have*

$$\left|\left\langle Q^{\kappa(m)}_{\Gamma,\mathsf{p}}, Q^{\kappa(n)}_{\Delta,\mathsf{r}}\right\rangle^\kappa_\Delta\right|^2 = w_\Gamma(\mathsf{r})\delta_{\Gamma,\mathsf{p}|\Gamma\mathsf{r}}\delta_{mw_\Gamma(\mathsf{r}),n} \tag{5.3.10}$$

for all $\mathsf{r} \in \mathsf{P}_\Delta$ and $n < 0$, for any choice of scaling coset systems for Δ and Γ.

To conclude this section we comment on the branching of Rademacher sums in the case that Γ both contains and normalizes Δ. Suppose then that Δ and Γ have width one at infinity, and consider the case that p is the infinite cusp $\Gamma\cdot\infty$ in (5.3.2). Then we are speaking of the map

$$\begin{aligned} \Delta\backslash\Gamma &\to \Delta\backslash\Gamma\cdot\infty \\ \Delta\gamma &\mapsto \Delta\gamma\cdot\infty, \end{aligned} \tag{5.3.11}$$

and this map is injective (and thus bijective) when Γ normalizes Δ. Indeed, since Δ is supposed to have width one at infinity we have $(\gamma^{-1}\Delta\gamma)_\infty = \Delta_\infty = B(\mathbb{Z})$, so that every coset $\Delta\gamma \in \Delta\backslash\Gamma$ is indeed a scaling coset for Δ at $\mathsf{r} = \Delta\gamma\cdot\infty$, and the injectivity of (5.3.11) follows from Lemma 5.3.1. Applying Theorem 5.3.4 now with $\mathsf{p} = \Gamma\cdot\infty$ we obtain the following result.

Proposition 5.3.6. *Suppose that Δ and Γ are groups commensurable with $G(\mathbb{Z})$ that both have width one at infinity, and suppose that Γ contains and normalizes Δ. Let $\kappa, m \in \mathbb{Z}$ such that*

$\kappa \leq 0$ and $m < 0$. Then we have

$$Q_{\Gamma}^{\kappa(m)}(z) = \sum_{\mathsf{r} \in \Delta \backslash \mathsf{p}} Q_{\Delta,\mathsf{r}}^{\kappa(m)}(z) \tag{5.3.12}$$

when the scaling cosets for Δ at the cusps $\mathsf{r} \in \Delta \backslash \Gamma \cdot \infty$ are taken to lie in $\Delta \backslash \Gamma$.

Theorem 5.3.5 now implies the following reformulation of Proposition 5.3.6 which is independent of scaling coset choices.

Proposition 5.3.7. *Suppose that Δ and Γ are groups commensurable with $G(\mathbb{Z})$ that both have width one at infinity, and suppose that Γ contains and normalizes Δ. Then we have*

$$\left| \left\langle Q_{\Gamma}^{\kappa(-1)}, Q_{\Delta,\mathsf{r}}^{\kappa(-1)} \right\rangle_{\Delta}^{\kappa} \right|^2 = \delta_{\Gamma,\mathsf{p}|\Gamma\mathsf{r}} \tag{5.3.13}$$

for all $\mathsf{r} \in \mathsf{P}_{\Delta}$, for any choice of scaling coset system for Δ.

5.4 Fractional orders

Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\phi, \psi \in G(\mathbb{Q})$ and set $U = [\![\phi^{-1}\Gamma\psi$. Let us consider the problem of writing the modified Rademacher sum $Q_U^{\kappa(m)}(z)$ in terms of Rademacher sums of the form $Q_{\Gamma',\mathsf{p}'}^{\kappa(m')}(z)$ for some group Γ' and some cusp $\mathsf{p}' \in \mathsf{P}_{\Gamma'}$. According to §4.1 we have $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m)}(z) = Q_{\Gamma',\mathsf{p}'}^{\kappa(m)}(z)$ for $\Gamma' = \Gamma^{\mathsf{q}}$ and $\mathsf{p}' = \mathsf{p}^{\mathsf{q}}$ (cf. (4.1.9)), so we are done if we can write $Q_U^{\kappa(m)}(z)$ in terms of $Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(n)}(z)$ for some cusps $\mathsf{p}, \mathsf{q} \in \mathsf{P}_{\Gamma}$ for Γ. Set $\mathsf{p} = \Gamma\phi \cdot \infty$ and $\mathsf{q} = \Gamma\psi \cdot \infty$. By the discussion of §2.6 there exist unique $\mu, \nu \in \mathbb{Q}^+$ such that the cosets $\Sigma_{\mathsf{p}} = \Gamma\phi[\mu]$ and $\Sigma_{\mathsf{q}} = \Gamma\psi[\nu]$ (cf. (2.3.3)) are scaling cosets for Γ at p and q, respectively. We then have $U = [\![[\mu]\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}[1/\nu]$. Let $\chi \in \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{q}}$ and consider the contribution $Q_{[\![[\mu]\chi[1/\nu]}^{\kappa(m)}(z,s)$ of the coset $[\![[\mu]\chi[1/\nu] \in U$ to the modified continued Rademacher sum $Q_U^{\kappa(m)}(z,s)$. As the following lemma demonstrates, the factor $[1/\nu]$ induces a re-scaling of the input variable z.

Lemma 5.4.1. *Let $\chi \in G(\mathbb{Q})$ and let $\nu \in \mathbb{Q}^+$. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then we have*

$$Q_{[\![\chi[1/\nu]}^{\kappa(m)}(z,s) = \frac{1}{\nu^{\kappa}} Q_{[\![\chi}^{\kappa(m)}\left(\frac{z}{\nu}, s\right). \tag{5.4.1}$$

Proof. We recall from (4.1.4) that the modified continued Rademacher component function $Q_{[\![\chi[1/\nu]}^{\kappa(m)}(z,s)$ is, by definition, the difference $T_{[\![\chi[1/\nu]}^{\kappa(m)}(z,s) - T_{0,[\![\chi[1/\nu]}^{\kappa(m)}(s)$, where the functions $T_{[\![\chi[1/\nu]}^{\kappa(m)}(z,s)$ and $T_{0,[\![\chi[1/\nu]}^{\kappa(m)}(s)$ are defined by (4.1.1) and (4.1.3), respectively. We compute $\operatorname{jac}([\![\chi[1/\nu]\,, z) = \operatorname{jac}([\![\chi\,, z/\nu)/\nu$, and use the fact that $[1/\nu]$ fixes both ∞ and 0 to verify that

$$T_{[\![\chi[1/\nu]}^{\kappa(m)}(z,s) = \frac{1}{\nu^{\kappa}} T_{[\![\chi}^{\kappa(m)}\left(\frac{z}{\nu}, s\right), \quad T_{0,[\![\chi[1/\nu]}^{\kappa(m)}(s) = \frac{1}{\nu^{\kappa}} T_{0,[\![\chi}^{\kappa(m)}(s). \tag{5.4.2}$$

The claim follows from these identities. □

We now seek to describe the effect of the factor $[\mu]$ in $Q^{\kappa(m)}_{[\![\mu]\chi}(z,s)$, for arbitrary $\chi \in G(\mathbb{Q})$. Inspecting (4.1.1) we find that

$$T^{\kappa(m)}_{[\![\mu]\chi}(z,s) = \mu^{\kappa}\,\mathrm{e}(m\mu\chi\cdot z)\,\mathrm{t}^{\kappa}(m\mu,\chi,z,s)\,\mathrm{jac}(\chi,z)^{\kappa}, \tag{5.4.3}$$

which suggests that we generalize the notion of Rademacher sum so as to allow for fractional orders. Recall from (2.3.5) that $B(\alpha\mathbb{Z})$ denotes the subgroup of $B_u(\mathbb{Q})$ generated by T^{α} (cf. (2.3.4)). For general $\mu \in \mathbb{Q}^+$ the right hand side of (5.4.3) will not be invariant under the replacement of χ by $T\chi$, so a Rademacher sum of fractional order is not naturally defined by collections of cosets of $B(\mathbb{Z})$, but rather, by cosets of $B(h\mathbb{Z})$ for a suitably chosen positive integer $h \in \mathbb{Z}^+$.

Given $h \in \mathbb{Z}^+$ and $\chi \in G(\mathbb{Q})$ let us write $\llbracket^h\chi$ as a shorthand for $B(h\mathbb{Z})\chi$, and for $X \subset G(\mathbb{Q})$ let us write $\llbracket^h X$ as a shorthand for the set of right cosets of $B(h\mathbb{Z})$ determined by elements of X, so that

$$\llbracket^h X = \{B(h\mathbb{Z})\chi \mid \chi \in X\}\,. \tag{5.4.4}$$

Let $\kappa \in \mathbb{Z}$ such that $\kappa \leq 0$. We define modified and normalized fractional Rademacher sums as follows, in analogy with the constructions of §4.1. For $\mu = -g/h$ with $(g,h) = 1$ and $g,h \in \mathbb{Z}^+$, and for $\llbracket^h\chi \in \llbracket^h G(\mathbb{Q})$, we define the *continued fractional Rademacher component function of weight* 2κ *and order* μ *associated to* $\llbracket^h\chi$, denoted $(z,s) \mapsto T^{\kappa(\mu)}_{\llbracket^h\chi}(z,s)$, by setting

$$T^{\kappa(\mu)}_{\llbracket^h\chi}(z,s) = \mathrm{e}(\mu\llbracket^h\chi\cdot z)\,\mathrm{t}^{\kappa}(\mu,\llbracket^h\chi,z,s)\,\mathrm{jac}(\llbracket^h\chi,z)^{\kappa}, \tag{5.4.5}$$

where $\mathrm{jac}(\llbracket^h\chi,z) = \mathrm{jac}(\chi,z)$, and $\mathrm{t}^{\kappa}(\mu,\llbracket^h\chi,z,s)$ is the *continued fractional Rademacher regularization factor of weight* 2κ, which is in turn given by

$$\mathrm{t}^{\kappa}(\mu,\llbracket^h\chi,z,s) = \Phi(s-2\kappa,1+s-2\kappa,\mu\llbracket^h\chi\cdot\infty-\mu\llbracket^h\chi\cdot z)(2\pi\mathbf{i}(\mu\llbracket^h\chi\cdot z-\mu\llbracket^h\chi\cdot\infty))^{s-2\kappa} \tag{5.4.6}$$

in case $\llbracket^h\chi \in \llbracket^h G(\mathbb{Q})^{\times}$, and $\mathrm{t}^{\kappa}(\mu,\llbracket^h\chi,z,s) = 1$ otherwise. We define a function $s \mapsto T^{\kappa(\mu)}_{0,\llbracket^h\chi}(s)$ by setting

$$\begin{aligned} T^{\kappa(\mu)}_{0,\llbracket^h\chi}(s) = {} & \mathrm{e}(\mu\llbracket^h\chi\cdot\infty)\left(2\pi\mathbf{i}(\mu\llbracket^h\chi\cdot 0-\mu\llbracket^h\chi\cdot\infty)\right)^{(s-2\kappa)}\mathrm{jac}(\llbracket^h\chi,0)^{\kappa} \\ & -\mathrm{e}(s/2)\,\mathrm{e}(\mu\llbracket^h\chi\cdot\infty)\left(2\pi\mathbf{i}(\mu\llbracket^h\chi\cdot\infty-\mu\llbracket^h\chi\cdot 0)\right)^{(s-2\kappa)}\mathrm{jac}(\llbracket^h\chi,0)^{\kappa} \end{aligned} \tag{5.4.7}$$

in case $\llbracket^h\chi \in \llbracket G(\mathbb{Q})^{\times\times}$, and by setting $T^{\kappa(\mu)}_{0,\llbracket^h\chi}(s) = 0$ otherwise, and we define the *modified* *continued fractional Rademacher component function of weight* 2κ *and order* μ *associated to* $\llbracket^h\chi$, denoted $(z,s)\mapsto Q^{\kappa(\mu)}_{\llbracket^h\chi}(z,s)$, by subtracting $T^{\kappa(\mu)}_{0,\llbracket^h\chi}(z,s)$ from $T^{\kappa(\mu)}_{\llbracket^h\chi}(z,s)$.

$$Q^{\kappa(\mu)}_{\llbracket^h\chi}(z,s)=T^{\kappa(\mu)}_{\llbracket^h\chi}(z,s)-T^{\kappa(\mu)}_{0,\llbracket^h\chi}(z,s) \tag{5.4.8}$$

For $U\subset\llbracket^hG(\mathbb{Q})$ we now define the *continued fractional Rademacher sum of weight* 2κ *and order* μ *associated to* U, denoted $T^{\kappa(\mu)}_U(z,s)$, and the *modified continued fractional Rademacher sum of weight* 2κ *and order* μ *associated to* U, denoted $Q^{\kappa(\mu)}_U(z,s)$, by setting

$$T^{\kappa(\mu)}_U(z,s)=\sum_{\llbracket^h\chi\in U}T^{\kappa(\mu)}_{\llbracket^h\chi}(z,s), \tag{5.4.9}$$

$$Q^{\kappa(\mu)}_U(z,s)=\sum_{\llbracket^h\chi\in U}Q^{\kappa(\mu)}_{\llbracket^h\chi}(z,s), \tag{5.4.10}$$

when these sums are absolutely locally uniformly convergent, and given such circumstances we define $T^{\kappa(\mu)}_U(z)$ and $Q^{\kappa(\mu)}_U(z)$ by taking the limit as s tends to 1 in $T^{\kappa(\mu)}_U(z,s)$ and $Q^{\kappa(\mu)}_U(z,s)$, respectively, so long as these limits exist.

$$T^{\kappa(\mu)}_U(z)=\lim_{s\to1^+}T^{\kappa(\mu)}_U(z,s) \tag{5.4.11}$$

$$Q^{\kappa(\mu)}_U(z)=\lim_{s\to1^+}Q^{\kappa(\mu)}_U(z,s) \tag{5.4.12}$$

We call $T^{\kappa(\mu)}_U(z)$ the *normalized fractional Rademacher sum of weight* 2κ *and order* μ *associated to* U, and we call $Q^{\kappa(\mu)}_U(z)$ the *modified fractional Rademacher sum of weight* 2κ *and order* μ *associated to* U.

We have the following vanishing result for fractional Rademacher sums.

Proposition 5.4.2. *Let* X *be a union of cosets of* $B(\mathbb{Z})$ *in* $G(\mathbb{Q})$, *let* $\mu=-g/h$ *for some* $g,h\in\mathbb{Z}^+$ *and* $(g,h)=1$, *and set* $U=\llbracket^hX$. *Suppose that the normalized fractional Rademacher sum* $T^{\kappa(\mu)}_U(z)$ *and the modified fractional Rademacher sum* $Q^{\kappa(\mu)}_U(z)$ *converge. Then they vanish identically unless* $h=1$.

Proof. Suppose that $h\neq1$. By hypothesis we have a disjoint decomposition $X=\bigcup_i\llbracket\chi_i$ for some $\{\chi_i\}\subset X$. This implies a disjoint decomposition $X=\bigcup_i\bigcup_{k=0}^{h-1}\llbracket^hT^k\chi_i$, so that we have

$$\begin{aligned}T^{\kappa(\mu)}_U(z,s)&=\sum_i\sum_{k=0}^{h-1}\mathrm{e}\left(-\frac{g}{h}T^k\chi_i\cdot z\right)\mathrm{t}^\kappa\left(-\frac{g}{h},T^k\chi_i,z,s\right)\mathrm{jac}(T^k\chi_i,z)^\kappa\\&=\sum_i\left(\sum_{k=0}^{h-1}\mathrm{e}\left(-\frac{gk}{h}\right)\right)\mathrm{e}\left(-\frac{g}{h}\chi_i\cdot z\right)\mathrm{t}^\kappa\left(-\frac{g}{h},\chi_i,z,s\right)\mathrm{jac}(\chi_i,z)^\kappa\end{aligned} \tag{5.4.13}$$

since the continued fractional regularization factor is unaffected when the second argument is multiplied by an element of $B(\mathbb{Z})$ on the left. The sum $\sum_{k=0}^{h-1} \mathrm{e}(-gk/h)$ vanishes for g coprime to h unless $h = 1$. This shows that the continued fractional Rademacher sum $T_U^{\kappa(\mu)}(z,s)$ vanishes for all z and s when $h \neq 1$. A directly analogous computation shows that the function $T_{0,U}^{\kappa(\mu)}(s)$ vanishes in case $h \neq 1$. We conclude that both the continued fractional Rademacher sum $T_U^{\kappa(\mu)}(z,s)$ and the modified continued fractional Rademacher sum $Q_U^{\kappa(\mu)}(z,s)$ vanish identically in case $h \neq 1$. The claim of the proposition follows. □

We return now to the identification of the Rademacher sum $Q_U^{\kappa(m)}(z)$, in the case that $U = \llbracket [\mu]\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}[1/\nu]$.

Theorem 5.4.3. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps for Γ, and let Σ_{p} and Σ_{q} be scaling cosets for Γ at p and q, respectively. Let $\mu,\nu \in \mathbb{Q}^+$ and set $U = \llbracket [\mu]\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}[1/\nu]$. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Let h be the smallest positive integer such that $h\mu \in \mathbb{Z}$. If h divides m then we have*

$$Q_U^{\kappa(m)}(z) = h\frac{\mu^\kappa}{\nu^\kappa} Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m\mu)}\left(\frac{z}{\nu}\right), \tag{5.4.14}$$

and if h does not divide m then $Q_U^{\kappa(m)}(z)$ vanishes identically.

Proof. Set $U' = \llbracket [\mu]\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$ and $U'' = \llbracket \Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}}$, so that $U = U'[1/\nu] = [\mu]U''[1/\nu]$. By Lemma 5.4.1 we have $Q_U^{\kappa(m)}(z) = Q_{U'}^{\kappa(m)}(z/\nu)/\nu^\kappa$, so it suffices for us to show that $Q_{U'}^{\kappa(m)}(z)$ is $h\mu^\kappa Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m\mu)}(z)$ or vanishing, according as h divides m or not.

Set $X = \Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{q}} \subset G(\mathbb{Q})$. Then X is a union of right cosets of $B(\mathbb{Z})$, so we have a disjoint decomposition $X = \bigcup_i \llbracket \chi_i$ for some $\chi_i \in G(\mathbb{Q})$. By the choice of h we have $\mu = g/h$ for some $g \in \mathbb{Z}^+$ with $(g,h) = 1$. Then for the set $[\mu]X$ we have

$$[\mu]X = \bigcup_i [\mu] \llbracket \chi_i = \bigcup_i \bigcup_{k=0}^{h-1} [\mu]T^k \llbracket^h \chi_i = \bigcup_i \bigcup_{k=0}^{h-1} T^{k\mu} \llbracket^g [\mu]\chi_i \tag{5.4.15}$$

with all the unions disjoint, from which we conclude that $\llbracket [\mu]X$ admits the disjoint decomposition $\llbracket [\mu]X = \bigcup_{k=0}^{h-1} \bigcup_i T^{k\mu} \llbracket [\mu]\chi_i$. We have

$$Q_{T^{k\mu}\llbracket [\mu]\chi_i}^{\kappa(m)}(z,s) = \mathrm{e}(mk\mu) Q_{\llbracket [\mu]\chi_i}^{\kappa(m)}(z,s). \tag{5.4.16}$$

If h divides m then $m\mu \in \mathbb{Z}$ and $\mathrm{e}(mk\mu) = 1$ for all k. Further, $Q_{\llbracket [\mu]\chi_i}^{\kappa(m)}(z,s)$ coincides with $\mu^\kappa Q_{\llbracket \chi_i}^{\kappa(m\mu)}(z,s)$ in this case, and so we deduce the required identity $Q_{U'}^{\kappa(m)}(z) = h\mu^\kappa Q_{\Gamma,\mathsf{p}|\mathsf{q}}^{\kappa(m\mu)}(z)$. On the other hand, if h does not divide m, so that $m\mu$ is not an integer, then the sum $\sum_{k=0}^{h-1} \mathrm{e}(mk\mu)$ vanishes and this implies the vanishing of the modified Rademacher sum $Q_{U'}^{\kappa(m)}(z)$. The proof is complete. □

Taking $\mathsf{q} = \Gamma \cdot \infty$ in Theorem 5.4.3 we obtain the following result.

Theorem 5.4.4. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let Z be a finite union of left cosets of Γ in $G(\mathbb{Q})$ and set $U = [\![Z$. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then the modified Rademacher sum $Q_U^{\kappa(m)}(z)$ is an automorphic integral of weight 2κ for Γ.*

5.5 Hecke operators

For the modified Rademacher sums, in contrast to the classical Rademacher sums, it is natural to consider not only their variance with respect to group actions, but also how they vary with respect to Hecke operators. The discrepancy in constant terms between the classical and modified Rademacher sums illustrated by Proposition 4.4.10 is a stubborn barrier to an exposition of the interaction between Hecke operators and the classical Rademacher sums.

Recall that $I_\kappa(\Gamma)$ denotes the space of automorphic integrals of weight 2κ for Γ, in the sense of §2.7, and recall the operator $I_\kappa(\Gamma) \to \mathcal{O}(\mathbb{H})$ of (2.7.14), denoted $f \mapsto f\|_\Gamma^\kappa X$, and defined for a right coset $X \in \Gamma\backslash G(\mathbb{Q})$ by setting $f\|_\Gamma^\kappa X = (f - J^\kappa_{\chi\cdot\infty} g)|^\kappa \chi$ for any representative $\chi \in X$ where $g \in S_{1-\kappa}(\Gamma)$ is the cusp form associated to f (cf. (2.7.13)). Let us generalize this operator by setting

$$f\|_\Gamma^\kappa X = \sum_i (f - J^\kappa_{\chi\cdot\infty} g)|^\kappa \chi_i \tag{5.5.1}$$

in case X is a finite union of right cosets of Γ in $G(\mathbb{Q})$ and the χ_i furnish a transversal $X = \bigcup_i \Gamma\chi_i$ for X over Γ. The next lemma verifies that the operator $f \mapsto f\|_\Gamma^\kappa X$ makes sense in case X is a double coset $X = \Gamma\sigma\Gamma$ of Γ in $G(\mathbb{Q})$.

Lemma 5.5.1. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let $\sigma \in G(\mathbb{Q})$, set $\Delta = \Gamma \cap \sigma^{-1}\Gamma\sigma$ and $\Delta' = \sigma\Gamma\sigma^{-1} \cap \Gamma$, and suppose that*

$$\Gamma = \bigcup_i \lambda_i'\Delta', \quad \Gamma = \bigcup_i \Delta\rho_i, \tag{5.5.2}$$

are left and right transversals for Γ over Δ' and Δ, respectively. Then the double coset $\Gamma\sigma\Gamma$ admits disjoint decompositions

$$\Gamma\sigma\Gamma = \bigcup_i \lambda_i'\sigma\Gamma, \quad \Gamma\sigma\Gamma = \bigcup_i \Gamma\sigma\rho_i, \tag{5.5.3}$$

into left and right cosets for Γ.

Proof. Since $G(\mathbb{Q})$ is the commensurator of $G(\mathbb{Z})$, the intersection $\Delta = \Gamma \cap \sigma^{-1}\Gamma\sigma$ has finite index in both Γ and $\sigma^{-1}\Gamma\sigma$, so there are only finitely many λ_i' and ρ_i. Observe next that

$\sigma\Delta = \Delta'\sigma = \sigma\Gamma \cap \Gamma\sigma$. Consequently, we have $\Delta'\sigma \subset \sigma\Gamma$ and $\sigma\Delta \subset \Gamma\sigma$, so that

$$\Gamma\sigma\Gamma = \bigcup \lambda_i'\Delta'\sigma\Gamma \subset \bigcup \lambda_i'\sigma\Gamma\Gamma = \bigcup \lambda_i'\sigma\Gamma, \tag{5.5.4}$$

$$\Gamma\sigma\Gamma = \bigcup \Gamma\sigma\Delta\rho_i \subset \bigcup \Gamma\Gamma\sigma\rho_i = \bigcup \Gamma\sigma\rho_i. \tag{5.5.5}$$

The reverse inclusions hold since all the λ_i' and ρ_i lie in Γ. The unions are disjoint, for if $\Gamma\sigma\rho_i = \Gamma\sigma\rho_j$ say, then $\rho_i\rho_j^{-1} \in \sigma^{-1}\Gamma\sigma$, but $\rho_i\rho_j^{-1} \in \Gamma$ by our choice of the ρ_i, so $\rho_i\rho_j^{-1} \in \Delta$, and this implies $i = j$. A similar argument applies to the cosets $\lambda_i'\sigma\Gamma$. □

The right hand identity of (5.5.3) shows that the operator $f \mapsto f\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma$ is well-defined for $f \in I_\kappa(\Gamma)$ and $\sigma \in G(\mathbb{Q})$. The left hand identity of (5.5.3) shows that the function $f\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma$ again lies in $I_\kappa(\Gamma)$. Indeed, if the cusp form associated to $f \in I_\kappa(\Gamma)$ is $g \in S_{1-\kappa}(\Gamma)$ then the cusp form associated to $f\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma$ is $g\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma$. We call the operator $I_\kappa(\Gamma) \to I_\kappa(\Gamma)$ given by $f \mapsto f\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma$ the *weight 2κ Hecke operator associated to σ*. In order to ease notation we set

$$f\dagger^{\kappa}_{\Gamma}\sigma = f\|^{\kappa}_{\Gamma}\Gamma\sigma\Gamma \tag{5.5.6}$$

for $\sigma \in G(\mathbb{Q})$. We have $f\dagger^{\kappa}_{\Gamma}\sigma = \sum_i f|^{\kappa}\chi_i$ when $\{\chi_i\} \subset G(\mathbb{Q})$ is a right transversal $\Gamma\sigma\Gamma = \bigcup_i \Gamma\chi_i$ for Γ in $\Gamma\sigma\Gamma$. If σ belongs to the normalizer of Γ then we have $\Gamma\sigma\Gamma = \Gamma\sigma$ and hence $f\dagger^{\kappa}_{\Gamma}\sigma = (f - J^{\kappa}_{\sigma\cdot\infty}g)|^{\kappa}\sigma$ for any $f \in I_\kappa(\Gamma)$. More generally, Lemma 5.5.1 yields for us the formula

$$f\dagger^{\kappa}_{\Gamma}\sigma = \sum_i (f - J^{\kappa}_{\sigma\rho_i\cdot\infty}g)|^{\kappa}(\sigma\rho_i) \tag{5.5.7}$$

in the case that $\{\rho_i\}$ is a right transversal for Γ over the intersection $\Delta = \Gamma \cap \Gamma^\sigma$.

The action of the operator $f \mapsto f\|^{\kappa}_{\Gamma}X$ on holomorphic Poincaré series can be described without reference to transversals. Indeed, if U is the set of right cosets of $B(\mathbb{Z})$ determined by a union of left cosets of some group Γ commensurable with $G(\mathbb{Z})$, and if X is a finite union $\bigcup_j \Gamma\chi_j$ say, of right cosets of Γ, then, taking $\kappa > 1$ to ensure absolute (and locally uniform) convergence, we have

$$P^{\kappa(m)}_{U}\|^{\kappa}_{\Gamma}X = \sum_j P^{\kappa(m)}_{U}|^{\kappa}\chi_j = \sum_j P^{\kappa(m)}_{U\chi_j} = P^{\kappa(m)}_{UX}, \tag{5.5.8}$$

so that the Poincaré series are stable under the action of the operators $f \mapsto f\|^{\kappa}_{\Gamma}X$. In particular, for the Hecke operator $f \mapsto f\dagger^{\kappa}_{\Gamma}\sigma$ we have $P^{\kappa(m)}_{U}\dagger^{\kappa}_{\Gamma}\sigma = P^{\kappa(m)}_{U\sigma\Gamma}$. This identity $P^{\kappa(m)}_{U}\|^{\kappa}_{\Gamma}X = P^{\kappa(m)}_{UX}$ extends naturally to the modified Rademacher sums.

Proposition 5.5.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$, let Z be a finite union of left cosets of Γ in $G(\mathbb{Q})$ and set $U = \llbracket Z$. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. Then for X a* *finite union of right cosets of* Γ *in* $G(\mathbb{Q})$ *we have*

$$Q_U^{\kappa(m)}\|_\Gamma^\kappa X = Q_{UX}^{\kappa(m)}. \tag{5.5.9}$$

Proof. The modified Rademacher sum $Q_U^{\kappa(m)}(z)$ is an automorphic integral of weight 2κ for Γ according to Theorem 5.4.4, so it has an associated cusp form $g \in S_{1-\kappa}(\Gamma)$ say. We may express X as a disjoint union $X = \bigcup_i \Gamma\chi_i$ for some finitely many $\chi_i \in G(\mathbb{Q})$. We then compute

$$Q_U^{\kappa(m)}\|_\Gamma^\kappa X = \sum_i (Q_U^{\kappa(m)} - J_{\chi_i\cdot\infty}^\kappa g)|^\kappa \chi_i = \sum_i Q_{U\chi_i}^{\kappa(m)} = Q_{UX}^{\kappa(m)}. \tag{5.5.10}$$

This proves the claim. □

In the case that Γ is the modular group $G(\mathbb{Z})$ we have the classical Hecke operators $\hat{T}(n)$, defined for $n \in \mathbb{Z}^+$ by setting

$$n^{1-\kappa}(\hat{T}(n)f)(z) = \sum_{\substack{ad=n\\ 0\leq b<d}} f\left(\frac{az+b}{d}\right)\frac{a^\kappa}{d^\kappa} \tag{5.5.11}$$

for $f \in M_\kappa(\Gamma)$. We deduce the relation between the operators $\hat{T}(n)$ and $f \mapsto f\dagger_\Gamma^\kappa\sigma$, for $\Gamma = G(\mathbb{Z})$, by observing the following coincidence of disjoint unions of cosets and double cosets of Γ.

$$\bigcup_{\substack{ad=n\\ 0\leq b<d}} \Gamma\begin{bmatrix} a & b\\ 0 & d\end{bmatrix} = \bigcup_{\substack{ad=n\\ d|a}} \Gamma\begin{bmatrix} a & 0\\ 0 & d\end{bmatrix}\Gamma \tag{5.5.12}$$

Both sides of (5.5.12) are decompositions of the image in $G(\mathbb{Q})$ of the set of 2×2 matrices with integral entries and determinant n. From these decompositions we deduce the following result.

Lemma 5.5.3. *For* $n \in \mathbb{Z}^+$ *and* $\kappa \in \mathbb{Z}$ *we have*

$$n^{1-\kappa}(\hat{T}(n)f) = \sum_{d\in\mathbb{Z}^+,\, d^2|n} f\dagger_\Gamma^\kappa[n/d^2] \tag{5.5.13}$$

for any $f \in I_\kappa(\Gamma)$.

As a partial converse to Lemma 5.5.3, observe that for $\Gamma = G(\mathbb{Z})$ the action of any Hecke operator $f \mapsto f\dagger_\Gamma^\kappa\sigma$ can be expressed in terms of the Hecke operator associated to a diagonal element $[\mu] \in B(\mathbb{Q})$. Indeed, if $\sigma \in B(\mathbb{Q})$ then, according to (5.5.12), we have $\Gamma\sigma\Gamma = \Gamma[\mu]\Gamma$ for some $\mu \in \mathbb{Q}^+$. In case σ does not lie in $B(\mathbb{Q})$ we have $\sigma\cdot\infty = p \neq \infty$ for some $p \in \mathbb{Q}$. Since the modular group acts transitively on $\hat{\mathbb{Q}}$ there is some $\sigma_p \in \Gamma$ with $\sigma_p\cdot\infty = p$, and then $\Gamma\sigma\Gamma = \Gamma\tilde{\sigma}\Gamma$ where $\tilde{\sigma} = \sigma_p^{-1}\sigma$ evidently lies in $B(\mathbb{Q})$.

Lemma 5.5.4. *Let $\Gamma = G(\mathbb{Z})$ and let $\sigma \in G(\mathbb{Q})$. Then there exists $\mu \in \mathbb{Q}^+$ such that*

$$f\dagger^{\kappa}_{\Gamma}\sigma = f\dagger^{\kappa}_{\Gamma}[\mu] \tag{5.5.14}$$

for all $\kappa \in \mathbb{Z}$ and all $f \in I_\kappa(\Gamma)$.

We now consider the action of the Hecke operators $\hat{T}(n)$ on the modified Rademacher sum $Q^{(-1)}_{\Gamma}(z) = Q^{(-1)}_{\llbracket\Gamma}(z)$ of weight 0 and order -1 associated to the modular group $\Gamma = G(\mathbb{Z})$. From Theorem 4.5.3 we have that $Q^{(-m)}_{\llbracket\Gamma}(z)$ is an automorphic integral of weight 0 (that is, an abelian integral) for Γ for $m > 0$. Since there are no non-zero cusp forms of weight 2 for the modular group, we see from Corollary 4.5.5 that $Q^{(-m)}_{\llbracket\Gamma}(z)$ is in fact a Γ-invariant function on $\mathbb{H}$ for all $m \in \mathbb{Z}^+$. Consequently we have $Q^{(-1)}_{\llbracket\Gamma}\dagger^{0}_{\Gamma}[n] = Q^{(-1)}_{\llbracket\Gamma[n]\Gamma}$ for $n \in \mathbb{Z}^+$, by Proposition 5.5.2 and the definition (5.5.6) of $f\dagger^{\kappa}_{\Gamma}\sigma$. Suppose that n is square-free. Then, according to Lemma 5.5.3, the action of the operator $f \mapsto f\dagger^{0}_{\Gamma}[n]$ on $M_0(\Gamma)$ coincides with that of $n\hat{T}(n)$. We anticipate an application of Lemma 5.5.1. For $\sigma = [n]$ we have $\Delta' = \sigma\Gamma\sigma^{-1} \cap \Gamma = \Gamma^0(n)$, and for a left transversal of $\Delta' = \Gamma^0(n)$ in Γ we may take

$$\Gamma = \bigcup_{e\|n} \bigcup_{k=0}^{n/e-1} T^k S T^e \Delta' \tag{5.5.15}$$

where the first union is over exact divisors of n (cf. §2.1). We deduce that

$$\Gamma[n]\Gamma = \bigcup_{e\|n} \bigcup_{k=0}^{n/e-1} T^k S T^e [n]\Gamma \tag{5.5.16}$$

by Lemma 5.5.1. Now $\llbracket T^k X = \llbracket X$ for any subset $X \subset G(\mathbb{Q})$ and any $k \in \mathbb{Z}$, so we find that

$$n(\hat{T}(n)Q^{(-1)}_{\llbracket\Gamma}) = Q^{(-1)}_{\llbracket\Gamma[n]\Gamma} = \sum_{e\|n} Q^{(-1)}_{\llbracket ST^e[n]\Gamma}. \tag{5.5.17}$$

Lemma 5.5.5. *Let e be an exact divisor of n. Then we have $\llbracket ST^e[n]\Gamma = \llbracket[n/e^2]\Gamma$ when Γ is the modular group $G(\mathbb{Z})$.*

Proof. We compute

$$ST^e[n] = \begin{bmatrix} 0 & -1 \\ n & e \end{bmatrix}, \quad ST^e[n]ST^{n/e}S = \begin{bmatrix} n/e & -1 \\ 0 & e \end{bmatrix}, \tag{5.5.18}$$

and observe that $T^k A T^l = [n/e^2]$ for $k, l \in \mathbb{Z}$ such that $ke + ln/e = 1$, where A is given by $A = ST^e[n]ST^{n/e}S$. ☐

We can now write

$$n(\hat{T}(n)Q^{(-1)}_{\llbracket\Gamma}) = \sum_{e\|n} Q^{(-1)}_{\llbracket[n/e^2]\Gamma} \tag{5.5.19}$$

in the case that n is square-free, and we can employ the methods of §5.4 to rewrite (5.5.19) in terms of the fractional Rademacher sums. Observe that $B(g\mathbb{Z})[g/h] = [g/h]B(h\mathbb{Z})$. We thus have

$$\llbracket[\mu]\chi = \bigcup_{k=0}^{g-1} T^k \llbracket^g[\mu]\chi = \bigcup_{k=0}^{g-1} T^k[\mu] \llbracket^h\chi \tag{5.5.20}$$

for any $\chi \in G(\mathbb{Q})$ when $\mu = g/h$ with $g, h \in \mathbb{Z}^+$ and $(g,h) = 1$. We rewrite the right hand side of (5.5.19), taking $g = n/h$ and $h = e$, as

$$\sum_{h\|n} Q^{(-1)}_{\llbracket[n/h^2]\Gamma} = \sum_{h\|n} Q^{(-1)}_{[n/h^2]\llbracket^h\Gamma} = \sum_{h\|n} Q^{(-n/h^2)}_{\llbracket^h\Gamma}, \tag{5.5.21}$$

and thus obtain the following result, relating the actions of Hecke operators to the fractional Rademacher sums.

Theorem 5.5.6. *Let $n \in \mathbb{Z}^+$ be square-free and let $\Gamma = G(\mathbb{Z})$. Then we have*

$$n(\hat{T}(n)Q^{(-1)}_{\llbracket\Gamma}) = \sum_{h\|n} Q^{(-n/h^2)}_{\llbracket^h\Gamma} \tag{5.5.22}$$

where the sum is over the exact divisors of n.

It is interesting to compare the result (5.5.22) of Theorem 5.5.6 with the identity

$$n\left(\hat{T}(n)Q^{(-1)}_{\llbracket\Gamma}\right)(z) = Q^{(-n)}_{\llbracket\Gamma}(z), \tag{5.5.23}$$

which holds since both sides are holomorphic on $\mathbb{H}$, and have the same singular terms in their Fourier expansion at $\Gamma \cdot \infty$. This proves that the Rademacher sum $Q^{(n/h^2)}_U(z)$ vanishes for $U = \llbracket^h\Gamma$ whenever h is an exact divisor of n, and thus recovers a special case of the vanishing result of Theorem 5.4.3.

6 Moonshine

Monstrous moonshine associates a group Γ_g commensurable with $G(\mathbb{Z})$ and having width one at infinity to each (conjugacy class of) element(s) g in the Monster group $\mathbb{M}$. In the article [CMS04] the set of groups $\{\Gamma_g \mid g \in \mathbb{M}\}$ is characterized in purely group theoretic terms. In this section we describe a reformulation of this characterization in terms of normalized Rademacher sums, and a certain family of moduli problems for solid tori with conformal structure on the boundary.

6.1 Genera

The main theorem of [CMS04] gives four conditions which, taken together, characterize the groups Γ_g for $g \in \mathbb{M}$. The first of these conditions is the requirement that a group Γ have genus zero. Our first result in this section is a reformulation of the genus zero condition in terms of Rademacher sums.

Theorem 6.1.1. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let p be a cusp for Γ. Then the normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ defines a Γ-invariant function on $\mathbb{H}$ if and only if Γ has genus zero.*

Proof. Let Γ be as in the statement of the theorem and let $\mathsf{p},\mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ. By Theorem 4.4.6 we have

$$T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z) = \delta_{\Gamma,\mathsf{p}|\mathsf{q}}\,\mathrm{e}(-z) + F^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)_{\mathrm{van}}. \tag{6.1.1}$$

Taking $\mathsf{q} = \Gamma \cdot \infty$ in (6.1.1) we see that $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is holomorphic on $\mathbb{H}$, and letting q range over P_Γ we see from (6.1.1) that the only pole of $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is a simple pole at the cusp p. We see then that if the function $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is Γ-invariant then it defines a morphism ϕ say, of Riemann surfaces $\phi : \mathsf{X}_\Gamma \to \hat{\mathbb{C}}$ (cf. §2.5) which has degree one since the preimage of $\infty \in \hat{\mathbb{C}}$ under ϕ is the single point $\mathsf{p} \in \mathsf{X}_\Gamma$ and ϕ is unramified here since Γ is assumed to have width one at infinity. We conclude that ϕ is an isomorphism, so that Γ indeed has genus zero.

Conversely, if Γ has genus zero then, since the space $S_1(\Gamma)$ of cusp forms of weight 2 for Γ is isomorphic to the space of holomorphic differentials on X_Γ, we have $S_1(\Gamma) = \{0\}$, so that any automorphic integral of weight 0 for Γ is in fact a Γ-invariant function, by Corollary 4.5.5. The modified Rademacher sum $Q^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is an automorphic integral for Γ by Theorem 4.5.3. The normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ differs from $Q^{(-1)}_{\Gamma,\mathsf{p}}(z)$ by a constant function by Proposition 4.4.11, and thus $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is Γ-invariant whenever $Q^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is. We conclude that $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ is Γ-invariant in case Γ has genus zero. This completes the proof. □

From the first part of the proof of Theorem 6.1.1 we see that the normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ associated to Γ at a cusp p defines an isomorphism $\mathsf{X}_\Gamma \to \hat{\mathbb{C}}$ mapping p to $\infty \in \hat{\mathbb{C}}$ in case Γ is a group of genus zero.

Theorem 6.1.2. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ. If Γ has genus zero then the normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}}(z)$ associated to Γ at the cusp p induces an isomorphism $\mathsf{X}_\Gamma \to \hat{\mathbb{C}}$ mapping p to ∞.*

Applying Theorem 6.1.2 with $\mathsf{p} = \Gamma \cdot \infty$ the infinite cusp we see that the normalized Rademacher sum $T_\Gamma^{(-1)}(z)$ induces an isomorphism $\mathsf{X}_\Gamma \to \hat{\mathbb{C}}$ mapping the infinite cusp to $\infty \in \hat{\mathbb{C}}$. Applying Proposition 4.4.9 with $\mathsf{p} = \mathsf{q} = \Gamma \cdot \infty$ we see that the Fourier expansion of $T_\Gamma^{(-1)}(z)$ has vanishing constant term. This shows that the normalized Rademacher sum $T_\Gamma^{(-1)}(z)$ associated to Γ at the infinite cusp is the normalized hauptmodul for Γ (cf. §2.5) whenever this statement makes sense; i.e. whenever Γ has genus zero.

Theorem 6.1.3. *Let Γ be a group commensurable with $G(\mathbb{Z})$ that has width one at infinity. If Γ has genus zero then the normalized Rademacher sum $T_\Gamma^{(-1)}(z)$ is the normalized hauptmodul for Γ.*

From the second part of the proof of Theorem 6.1.1 we see that if Γ has genus zero then not only the first order but in fact all the higher order normalized Rademacher sums $T_{\Gamma,\mathsf{p}}^{(-m)}(z)$ (for $m > 0$) associated to Γ at an arbitrary cusp $\mathsf{p} \in \mathsf{P}_\Gamma$ are Γ-invariant.

Theorem 6.1.4. *Let Γ be a subgroup of $G(\mathbb{Q})$ that is commensurable with $G(\mathbb{Z})$ and has width one at infinity and let $\mathsf{p} \in \mathsf{P}_\Gamma$ be a cusp of Γ. Let $m \in \mathbb{Z}^+$. If Γ has genus zero then the normalized Rademacher sum $T_{\Gamma,\mathsf{p}}^{(-m)}(z)$ is Γ-invariant.*

6.2 Moduli

The second condition of the main theorem of [CMS04] is that a group Γ be of $n\|h$-type (cf. §2.5). In this section we provide a geometric description of the groups of $n\|h$-type by furnishing a family of moduli problems for which the corresponding moduli spaces may be realized as quotients of the form $\mathsf{Y}_\Gamma = \Gamma\backslash\mathbb{H}$ where Γ is a group of $n\|h$-type.

Consider pairs (E, C) where E is an elliptic curve over $\mathbb{C}$ and C is an oriented subgroup of E isomorphic to S^1. We call such a pair a *solid torus*. For (E, C) a solid torus, we call E the *underlying elliptic curve*, and we call C the *underlying primitive cycle*. Note that for (E, C) a solid torus, the group C determines a primitive element of the first homology group of (the smooth real surface underlying) the elliptic curve E.

Given $z \in \mathbb{H}$ let us write Λ_z for the lattice $\mathbb{Z}z + \mathbb{Z} \subset \mathbb{C}$. Observe that the lattices Λ_z and $\Lambda_{z'}$ coincide if any only if $z - z' \in \mathbb{Z}$, so the assignment $z \mapsto \Lambda_z$ descends naturally to the orbit space $B(\mathbb{Z})\backslash\mathbb{H}$. If we agree to write $[\![z$ as a shorthand for the $B(\mathbb{Z})$ orbit determined by $z \in \mathbb{H}$, then we may unambiguously write $\Lambda_{[\![z}$ for the lattice $\mathbb{Z}z + \mathbb{Z}$. We write $E_{[\![z}$ for the corresponding elliptic curve $E_{[\![z} = \mathbb{C}/\Lambda_{[\![z} = \mathbb{C}/\mathbb{Z}z + \mathbb{Z}$. Then the pair $(E_{[\![z}, C_{[\![z})$ is a solid torus when we take the subgroup $C_{[\![z}$ to be the image of $\mathbb{R}$ in $E_{[\![z}$ under the composition $\mathbb{R} \to \mathbb{C} \to E_{[\![z}$. Observe that for any solid torus (E, C) there exists $z \in \mathbb{H}$ and a unique isomorphism of elliptic curves $E \to E_{[\![z}$ inducing an orientation preserving isomorphism $C \to C_{[\![z}$.

For (E, C) a solid torus and $n \in \mathbb{Z}^+$ we write $E[n]$ for the group of n-division points of E, and we write $C[n]$ for the intersection $C \cap E[n]$. Since C is oriented, each group $C[n]$ comes equipped with a distinguished generator; viz., the point corresponding to $1/n + \Lambda_{\llbracket z}$ under an isomorphism $E \to E_{\llbracket z}$ inducing an orientation preserving isomorphism $C \to C_{\llbracket z}$. If K is a subgroup of $E[n]$ for some n then the quotient $\bar{E} = E/K$ is again an elliptic curve, and the image $\bar{C}$ say, of C under the natural map $E \to \bar{E}$ defines a primitive cycle of $\bar{E}$, so that the pair $(\bar{E}, \bar{C})$ is again a solid torus. We call this the *quotient of* (E, C) *by* K and denote it $(E, C)/K$.

For $n \in \mathbb{Z}^+$ define an *n-compatible isogeny of solid tori* $(E', C') \to (E, C)$ to be an isogeny $E' \to E$ of elliptic curves that maps $C'[n]$ to a subgroup of $C[n]$. Then a 1-compatible isogeny of solid tori is just an isogeny of the underlying elliptic curves. We define an *isogeny of solid tori* $(E', C') \to (E, C)$ to be an isogeny $E' \to E$ of elliptic curves that restricts to an orientation preserving map $C' \to C$ on the underlying primitive cycles. For (E, C) a solid torus and $n \in \mathbb{Z}^+$, we may now interpret the canonical map $E \to E/C[n]$ as defining an isogeny $(E, C) \to (E, C)/C[n]$ of solid tori. We call this isogeny the *canonical n-fold quotient of* (E, C).

Observe that the elliptic curve $E/C[n]$ underlying the quotient $(E, C)/C[n]$ both receives a natural map from E and maps naturally to E, for the quotient $E/E[n]$ is naturally isomorphic to E, and for the map $E/C[n] \to E$ we may take the composition $E/C[n] \to E/E[n] \xrightarrow{\sim} E$ where the first map is the natural projection and the second map is the isomorphism just mentioned. This map $E/C[n] \to E$ defines an isogeny $(E, C)/C[n] \to (E, C)$ of solid tori; indeed, it restricts to an isomorphism on the underlying primitive cycles. We call the isogeny $(E, C)/C[n] \to (E, C)$ of solid tori the *canonical n-fold cover of* (E, C).

Say an n-compatible isogeny of solid tori $(E', C') \to (E, C)$, for $n \in \mathbb{Z}^+$, is an *n-compatible isomorphism of solid tori* in case it admits an inverse n-compatible isogeny; i.e. an isogeny $E \to E'$ mapping $C[n]$ to a subgroup of $C'[n]$ such that the compositions $E' \to E \to E'$ and $E \to E' \to E$ are the identity maps on E' and E, respectively. Simply put, an n-compatible isomorphism of solid tori $(E', C') \to (E, C)$ is an isomorphism of the underlying elliptic curves $E' \to E$ that induces an isomorphism of groups $C'[n] \to C[n]$. An *isomorphism of solid tori* $(E', C') \to (E, C)$ is an isomorphism $E' \to E$ of elliptic curves that restricts to an orientation preserving isomorphism $C' \to C$ on the underlying primitive cycles.

Suppose (E, C) and (E', C') are solid tori. For n a positive integer and e an exact divisor of n, say (E, C) and (E', C') are *$n+e$-related* if there is an n-compatible isogeny $(E', C') \to (E, C)$ that induces n/e-compatible isomorphisms $(E', C')/C'[e] \to (E, C)$ and $(E', C') \to (E, C)/C[e]$. More precisely, we require that the n-compatible isogeny $(E', C') \to (E, C)$ factor through the canonical e-fold maps $(E', C') \to (E', C')/C'[e]$ and $(E, C)/C[e] \to (E, C)$, defining n/e-

compatible isomorphisms of the specified type.

$$\begin{tikzcd}
(E',C') \arrow[r, "\sim", "n/e"'] \arrow[d] & (E,C)/C[e] \arrow[d] \\
(E',C')/C'[e] \arrow[r, "\sim", "n/e"'] & (E,C)
\end{tikzcd} \tag{6.2.1}$$

Then solid tori (E,C) and (E',C') are $n+1$-related just if they are n-compatible isomorphic, which is the case just if there is an isomorphism of elliptic curves $E' \to E$ mapping $C'[n]$ onto $C[n]$. In particular, the notion of being $n+1$-related is an equivalence relation on solid tori. For S a subset of the set of exact divisors of n, say solid tori (E,C) and (E',C') are *$n+S$-related* if (E,C) and (E',C') are $n+e$-related for some $e \in S$.

Recall from §2.2 that the set of exact divisors of n, denoted $\mathrm{Ex}(n)$, admits a naturally defined group structure.

Lemma 6.2.1. *Let $n \in \mathbb{Z}^+$ and let $S \subset \mathrm{Ex}(n)$. Then the $n+S$-relation is an equivalence relation on solid tori if and only if S is a subgroup of* $\mathrm{Ex}(n)$.

In light of Lemma 6.2.1, we say that solid tori (E,C) and (E',C') are *$n+S$-equivalent* if they are $n+S$-related and S is a subgroup of the group of exact divisors of n.

In a slight modification of the notation of [CN79], we write $\Gamma_0(n)+S$ for the group formed by taking the union of the Hecke congruence group $\Gamma_0(n)$ and the Atkin–Lehner involutions $W_e(n)$ for e in S, when S is a subgroup of $\mathrm{Ex}(n)$.

$$W_e(n) = \left\{ \begin{bmatrix} ae & b \\ cn & de \end{bmatrix} \in G(\mathbb{Q}) \mid a,b,c,d \in \mathbb{Z},\, ade - bcn/e = 1 \right\} \tag{6.2.2}$$

Let n be a positive integer and let h be a divisor of n. Then an n-compatible isomorphism $\phi : (E',C') \to (E,C)$ naturally determines an nh-compatible isogeny $\hat{\phi} : (E',C') \to (E,C)$ which factors through an n/h-compatible isomorphism $\bar{\phi} : (E',C')/C'[h] \to (E,C)/C[h]$ on the canonical h-fold quotients, via the canonical maps $(E',C') \to (E',C')/C'[h]$ and $(E,C)/C[h] \to (E,C)$, as depicted in (6.2.3).

$$\begin{tikzcd}
(E',C') \arrow[r, "\hat{\phi}", "nh"'] \arrow[d] & (E,C) \\
(E',C')/C'[h] \arrow[r, "\bar{\phi}", "n/h"'] & (E,C)/C[h] \arrow[u]
\end{tikzcd} \tag{6.2.3}$$

We may take this nh-compatible isogeny $\hat{\phi}$ to be the composition $E' \to E \to E/E[h] \to E$ where the first map is the given map ϕ, the second map is the natural projection, and the third map is the natural isomorphism. Since ϕ is an n-compatible isomorphism the composition $E' \to E \to E/E[h]$ factors through the natural projection $E/C[h] \to E/E[h]$, and the kernel of the induced map $E' \to E/C[h]$ is just $C'[h]$, and so we arrive at an isomorphism $E'/C'[h] \to E/C[h]$ of elliptic curves, which defines the required n/h-compatible isomorphism $\bar{\phi} : (E', C')/C'[h] \to (E, C)/C[h]$ of the corresponding solid tori.

$$\begin{tikzcd}
(E', C') \arrow[r, "\phi", "n"'] \arrow[dr] & (E, C) \arrow[r] & (E, C)/E[h] \arrow[r, "\sim"] & (E, C) \\
& (E', C')/C'[h] \arrow[r, "\bar{\phi}", "n/h"'] & (E, C)/C[h] \arrow[u] \arrow[ur] &
\end{tikzcd} \tag{6.2.4}$$

More generally, and by the same argument, an n-compatible isogeny $(E', C') \to (E, C)$ naturally defines an nh-compatible isogeny $(E', C') \to (E, C)$ which factors through an n/h-compatible isogeny $(E', C')/C'[h] \to (E, C)/C[h]$, via the canonical maps $(E', C') \to (E', C')/C'[h]$ and $(E, C)/C[h] \to (E, C)$.

Not every n/h-compatible isogeny $\bar{\phi} : (E', C')/C'[h] \to (E, C)/C[h]$ arises in this way from an n-compatible isogeny $\phi : (E', C') \to (E, C)$, but, given an n/h-compatible isogeny $\bar{\phi} : (E', C')/C'[h] \to (E, C)/C[h]$, we always have the nh-compatible isogeny $\hat{\phi} : (E', C') \to (E, C)$ obtained as the composition

$$(E', C') \to (E', C')/C'[h] \to (E, C)/C[h] \to (E, C) \tag{6.2.5}$$

where the first map is the canonical h-fold quotient, and the third map is the canonical h-fold cover. We call this composition $\hat{\phi}$ the *nh-compatible isogeny induced from $\bar{\phi}$*.

For $n \in \mathbb{Z}^+$ and h a divisor of n say solid tori (E, C) and (E', C') are *$n|h$-related* if the canonical h-fold quotients $(E, C)/C[h]$ and $(E', C')/C'[h]$ are n/h-compatible isomorphic. For e an exact divisor of n/h, say (E, C) and (E', C') are *$n|h + e$-related* if the quotients $(E, C)/C[h]$ and $(E', C')/C'[h]$ are $n/h + e$-related. For S a subset of $\mathrm{Ex}(n/h)$ say (E, C) and (E', C') are *$n|h + S$-related* if they are $n|h + e$-related for some $e \in S$.

Lemma 6.2.2. *Let $n \in \mathbb{Z}^+$, let h be a divisor of n, and let $S \subset \mathrm{Ex}(n/h)$. Then the $n|h + S$-relation is an equivalence relation on solid tori if and only if S is a subgroup of $\mathrm{Ex}(n/h)$.*

Let N be a positive integer, set $h = ((N, 24)$ (cf. §2.2), and set $n = N/h$. Then for S a subgroup of the group of exact divisors of n/h we arrive at the notion of $n\|h + S$-equivalence of solid tori in the following way. Suppose (E, C) and (E', C') are $\Gamma_0(n|h) + S$-equivalent. Then we have isomorphisms $(E,C)\to(E_{\llbracket z},C_{\llbracket z})$ and $(E',C')\to(E_{\llbracket z'},C_{\llbracket z'})$ for some $z,z'\in\mathbb{H}$. Let S' be the image of S under the natural injection $\mathrm{Ex}(n/h)\to\mathrm{Ex}(N)$ (cf. §2.2), and define elements x, y and $w_{e'}$ for $e'\in S'$ by setting

$$x=\begin{bmatrix}1&1/h\\0&1\end{bmatrix},\quad y=\begin{bmatrix}1&0\\n&1\end{bmatrix},\quad w_{e'}=\begin{bmatrix}ae'&b\\cN&de'\end{bmatrix},\tag{6.2.6}$$

where the a, b, c and d in the definition of $w_{e'}$ are arbitrary integers for which we have $ade'-bcN/e'=1$. Then there is a sequence $(\gamma_0,\gamma_1,\cdots,\gamma_k)$ with $\gamma_0\in\Gamma_0(N)$ and $\gamma_i\in\{x,y,w_{e'}\mid e'\in S'\}$ for $1\le i\le k$ for which we have $z'=\gamma_k\cdots\gamma_1\gamma_0\cdot z$. We say now that (E,C) and (E',C') are $\Gamma_0(n\|h)+S$*-equivalent* in the case that the sum $\sum_{i=1}^k\lambda(\gamma_i)$ vanishes in $\mathbb{Z}/h$, where $\lambda(x)=1$, and $\lambda(w_{e'})=0$ for all $e'\in S'$, and $\lambda(y)$ is -1 or 1 according as $\Gamma_0(n|h)+S$ is Fricke or not.

Theorem 6.2.3. *Let $N\in\mathbb{Z}^+$, set $h=((N,24)$ and set $n=N/h$. Let S be a subgroup of $\mathrm{Ex}(n/h)$, and set $\Gamma=\Gamma_0(n\|h)+S$. Then the quotient $\Gamma\backslash\mathbb{H}$ is in natural correspondence with $n\|h+S$-equivalence classes of solid tori.*

6.3 Exponents

The third of the four conditions of the main theorem of [CMS04] states that if Γ is of the form $\Gamma_0(n\|h)+S$ then the quotient group $\Gamma/\Gamma_0(nh)$ should be a group of exponent two. In the present section we investigate the relationship between this exponent two condition and properties of scaling cosets.

Lemma 6.3.1. *Suppose that Δ is a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{r}\in\mathsf{P}_\Delta$ be a cusp of Δ. Let $\kappa,m\in\mathbb{Z}$ such that $\kappa\le0$ and $m<0$. If Σ_r is a scaling coset for Δ at r with the property that $\Sigma_\mathsf{r}^2=\Delta$ then*

$$Q^{\kappa(m)}_{\Delta,\mathsf{r}}(z)=Q^{\kappa(m)}_{\Delta|\mathsf{r}}(z),\tag{6.3.1}$$

and Σ_r is contained in the normalizer of Δ.

Proof. Set $U=\llbracket\Sigma_\mathsf{r}^{-1}$ and set $U'=\llbracket\Sigma_\mathsf{r}$. Then, under the assumption that Δ has width one at infinity, we have $Q^{\kappa(m)}_{\Delta,\mathsf{r}}(z)=Q^{\kappa(m)}_{U}(z)$ and $Q^{\kappa(m)}_{\Delta|\mathsf{r}}(z)=Q^{\kappa(m)}_{U'}(z)$. Let $\sigma\in\Sigma_\mathsf{r}$. Then $\Sigma_\mathsf{r}^2=\Delta$ implies $\sigma\Delta\sigma=\Delta$, which is equivalent to the identity $\Delta\sigma=\sigma^{-1}\Delta$. We conclude that $\Sigma_\mathsf{r}^{-1}=\Sigma_\mathsf{r}$, so that $U=U'$, so that $Q^{\kappa(m)}_{\Delta,\mathsf{r}}(z)=Q^{\kappa(m)}_{\Delta|\mathsf{r}}(z)$. The identity $\sigma\Delta\sigma=\Delta$ also implies $\sigma^2\in\Delta$, so that $\sigma\Delta\sigma=\sigma^{-1}\Delta\sigma$. We conclude that σ normalizes Δ, so that $\Sigma_\mathsf{r}=\Delta\sigma\in\Delta\backslash N(\Delta)$. This completes the proof. □

Lemma 6.3.1 states that if a scaling coset at a cusp r of a group Δ can be chosen so that its square is the trivial coset then we have a kind of commutativity for the Rademacher sums associated to Δ; viz. the expansion at infinity of the modified Rademacher sum associated to Δ and the cusp r coincides with the expansion at r of the modified Rademacher sum associated to Δ and the cusp at infinity.

We also have the following converse to Lemma 6.3.1.

Proposition 6.3.2. *Suppose that Δ is a group commensurable with $G(\mathbb{Z})$ that has width one at infinity and let $\mathsf{r} \in \mathsf{P}_\Delta$ be a cusp of Δ. Let $\kappa, m \in \mathbb{Z}$ such that $\kappa \leq 0$ and $m < 0$. If Σ_r is a scaling coset for Δ at r that is contained in the normalizer of Δ and*

$$Q^{\kappa(m)}_{\Delta,\mathsf{r}}(z) = Q^{\kappa(m)}_{\Delta|\mathsf{r}}(z) \tag{6.3.2}$$

then $\Sigma_\mathsf{r}^2 = \Delta$.

Proof. As in the proof of Lemma 6.3.1 we set $U = \llbracket \Sigma_\mathsf{r}^{-1}$ and $U' = \llbracket \Sigma_\mathsf{r}$. Then since Δ is assumed to have width one at infinity we have $Q^{\kappa(m)}_{\Delta,\mathsf{r}}(z) = Q^{\kappa(m)}_{U}(z)$ and $Q^{\kappa(m)}_{\Delta|\mathsf{r}}(z) = Q^{\kappa(m)}_{U'}(z)$. Let $\sigma \in \Sigma_\mathsf{r}$. Then we have $U = \llbracket \sigma^{-1}\Delta$ and $U' = \llbracket \Delta\sigma = \llbracket \sigma\Delta$. Set $\mathsf{r}' = \Delta\sigma^{-1} \cdot \infty$. Then $Q^{\kappa(m)}_{U'}(z) = Q^{\kappa(m)}_{\Delta,\mathsf{r}'}(z)$, so the coincidence $Q^{\kappa(m)}_{U}(z) = Q^{\kappa(m)}_{U'}(z)$ implies that $\mathsf{r} = \mathsf{r}'$. This in turn implies $\sigma^2 \in \Delta$, so that $\Sigma_\mathsf{r}^2 = \Delta$, as we required to show. ☐

6.4 Cusps

The forth condition of the main theorem of [CMS04] is the following. Supposing that a group Γ is of the form $\Gamma = \Gamma_0(n\|h) + S$ for some subgroup S of the group of exact divisors of n/h, for each $p \in \hat{\mathbb{Q}}$ there should exist an element $\tilde{\sigma}_p \in G(\mathbb{R})$ such that

$$\infty = \tilde{\sigma}_p \cdot p, \quad (\tilde{\sigma}_p \Gamma \tilde{\sigma}_p^{-1})_\infty = B(\mathbb{Z}), \quad \tilde{\sigma}_p \Gamma \tilde{\sigma}_p^{-1} \supset \Gamma_0(nh). \tag{6.4.1}$$

Observe that if we set $\sigma_p = \tilde{\sigma}_p^{-1}$, then the conditions of (6.4.1) translate into

$$\sigma_p \cdot \infty = p, \quad (\sigma_p^{-1} \Gamma \sigma_p)_\infty = B(\mathbb{Z}), \quad \sigma_p^{-1} \Gamma \sigma_p \supset \Gamma_0(nh), \tag{6.4.2}$$

the first two of which are just the conditions (2.6.1) of Lemma 2.6.1 except that we allow σ_p to lie in $G(\mathbb{R})$ in (6.4.2) but insist that σ_p belong to $G(\mathbb{Q})$ in (2.6.1). Actually, an element $\sigma_p \in G(\mathbb{R})$ satisfying (6.4.2) must lie in $G(\mathbb{Q})$, as the following result demonstrates.

Lemma 6.4.1. *Let Δ and Γ be groups commensurable with $G(\mathbb{Z})$ and suppose that Δ has width one at infinity. Let $p \in \hat{\mathbb{Q}}$ and suppose that $\sigma \in G(\mathbb{R})$ satisfies $\sigma \cdot \infty = p$ and $\sigma^{-1}\Gamma\sigma \supset \Delta$. Then $\sigma \in G(\mathbb{Q})$.*

Proof. Suppose that $\sigma \in G(\mathbb{R})$ satisfies the hypotheses of the lemma. Then $(\sigma^{-1}\Gamma\sigma)_\infty$ is an infinite cyclic group containing Δ_∞ and $\Delta_\infty = B(\mathbb{Z})$ since Δ has width one at infinity. It must be then that $(\sigma^{-1}\Gamma\sigma)_\infty$ is generated by $T^{1/n}$ for some $n \in \mathbb{Z}^+$, so that $([n]\sigma^{-1}\Gamma\sigma[1/n])_\infty = B(\mathbb{Z})$. According to Lemma 2.6.1 there exists $\sigma' \in G(\mathbb{Q})$ such that $\sigma'\cdot\infty = p$ and $((\sigma')^{-1}\Gamma\sigma')_\infty = B(\mathbb{Z})$. Now $[n]\sigma^{-1}\sigma'$ fixes ∞ and so lies in $B(\mathbb{R})$. Since σ' and $\sigma[1/n]$ both conjugate Γ to a group with width one at infinity we must in fact have $[n]\sigma^{-1}\sigma' \in B_u(\mathbb{R})$, where $B_u(\mathbb{R})$ consists of all the elements T^α (cf. (2.3.4)) with $\alpha \in \mathbb{R}$. We conclude that $\sigma = \sigma'' T^\alpha$ for some $\sigma'' \in G(\mathbb{Q})$ and $\alpha \in \mathbb{R}$. Let $\gamma \in \Gamma$ such that $\sigma^{-1}\gamma\sigma \in \Delta^\times$, so that $\sigma^{-1}\gamma\sigma\cdot\infty = q$ for some $q \in \mathbb{Q}$. Then $q = q'' - \alpha$ for $q'' = (\sigma'')^{-1}\gamma\sigma''\cdot\infty$. Since both q and q'' lie in $\mathbb{Q}$ we conclude that α also lies in $\mathbb{Q}$. Then the identity $\sigma = \sigma''T^\alpha$ implies that $\sigma \in G(\mathbb{Q})$, as we required to show. ☐

Recall from §2.6 that for Γ a group commensurable with $G(\mathbb{Z})$, for $\mathsf{p} \in \mathsf{P}_\Gamma$ a cusp of Γ, and for $\Sigma_{\mathsf{p}} \in \Gamma\backslash G(\mathbb{Q})$ a scaling coset for Γ at p, we write Γ^{p} as a shorthand for the group $\Sigma_{\mathsf{p}}^{-1}\Sigma_{\mathsf{p}}$. In light of Lemma 6.4.1 we may reformulate the fourth condition of the main theorem of [CMS04] as follows.

Lemma 6.4.2. *Let Γ be a group of $n\|h$-type, so that $\Gamma = \Gamma_0(n\|h)+S$ for some positive integers n and h, and some subgroup S of the group of exact divisors of n/h. Then the fourth condition of the main theorem of [CMS04] is satisfied if and only if for each cusp $\mathsf{p} \in \mathsf{P}_\Gamma$ of Γ there exists a scaling coset $\Sigma_{\mathsf{p}} \in \Gamma\backslash G(\mathbb{Q})$ for Γ at p such that Γ^{p} contains $\Gamma_0(nh)$.*

We conclude this section by relating the reformulation of Lemma 6.4.2 to the normalized Rademacher sums associated to genus zero groups of $n\|h$-type.

Proposition 6.4.3. *Let Γ be a group of $n\|h$-type, so that $\Gamma = \Gamma_0(n\|h) + S$ for some positive integers n and h, and some subgroup S of the group of exact divisors of n/h, and suppose that Γ has genus zero. Let $\mathsf{p}, \mathsf{q} \in \mathsf{P}_\Gamma$ be cusps of Γ and let Σ_{p} and Σ_{q} be scaling cosets for Γ at p and q, respectively. Then the normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is $\Gamma_0(nh)$-invariant if and only if Γ^{q} contains $\Gamma_0(nh)$.*

Proof. Let us set $\Delta = \Gamma_0(nh)$. Observe that if we set $\mathsf{p}^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}\cdot\infty$ then p^{q} is a cusp of $\Gamma^{\mathsf{q}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{q}}$, and $\Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}$ is a scaling coset for Γ^{q} at p^{q}, so the function $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ may be identified with the normalized Rademacher sum $T^{(-1)}_{\Gamma^{\mathsf{q}},\mathsf{p}^{\mathsf{q}}}(z)$ associated to Γ^{q} at the cusp p^{q}, for $\Sigma_{\mathsf{p}^{\mathsf{q}}} = \Sigma_{\mathsf{q}}^{-1}\Sigma_{\mathsf{p}}$. Under the assumption that Γ has genus zero, Γ^{q} also has genus zero, and has width one at infinity by the defining properties (cf. Lemma 2.6.2) of Σ_{q}. By Theorem 6.1.3 then the normalized Rademacher sum $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is, up to a constant function, the expansion at p^{q} of the normalized hauptmodul of the group Γ^{q}. In particular, $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ defines an isomorphism of Riemann surfaces $\mathsf{X}_{\Gamma^{\mathsf{q}}} \to \hat{\mathbb{C}}$, and so we have $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(\gamma\cdot z) = T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ for all $z \in \mathbb{H}$ if and only if $\gamma \in \Gamma^{\mathsf{q}}$. It follows then that $T^{(-1)}_{\Gamma,\mathsf{p}|\mathsf{q}}(z)$ is Δ-invariant if and only if Δ is contained in Γ^{q}, which is what we required to show. ☐

6.5 Rademacher sums and the Monster

We will now use Rademacher sums to reformulate the characterization of the groups of monstrous moonshine due to Conway–McKay–Sebbar (cf. [CMS04]).

Theorem 6.5.1. *Let Γ be a subgroup of $G(\mathbb{R})$. Then we have $\Gamma = \Gamma_g$ for some $g \in \mathbb{M}$ if and only if the following conditions are satisfied:*

- *Γ is the group defining $n\|h + S$-equivalence of solid tori for some positive integers n and h, and some subgroup $S < \mathrm{Ex}(n/h)$;*
- *the normalized Rademacher sum $T_{\Gamma}^{(-1)}(z)$ is Γ-invariant;*
- *there exists a system $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_{\Gamma}\}$ of scaling cosets for Γ such that $T_{\Gamma|\mathsf{p}}^{(-1)}(z)$ is Δ-invariant for every cusp $\mathsf{p} \in \mathsf{P}_{\Gamma}$ of Γ, where $\Delta = \Gamma_0(nh)$;*
- *for every cusp $\mathsf{r} \in \mathsf{P}_{\Delta}$ of $\Delta = \Gamma_0(nh)$ that is contained in $\Gamma\cdot\infty$ we have $T_{\Delta,\mathsf{r}}^{(-1)}(z) = T_{\Delta|\mathsf{r}}^{(-1)}(z)$ when the scaling coset Σ_{r} for Δ at r is taken to lie in $\Delta\backslash\Gamma$.*

Proof. According to the main theorem of [CMS04] it suffices to show that a group Γ satisfies each of the four conditions of Theorem 6.5.1 if and only if it satisfies the four conditions of the main theorem of [CMS04]. As a first step in establishing this equivalence, observe that the first condition of Theorem 6.5.1 is exactly the same as the second condition of the main theorem of [CMS04].

Suppose that Γ satisfies the four conditions of the main theorem of [CMS04]. Then, in particular, it satisfies the first condition of Theorem 6.5.1, and has genus zero. We may suppose then that $\Gamma = \Gamma_0(n\|h) + S$ for some $n, h \in \mathbb{Z}^+$, and $S < \mathrm{Ex}(n/h)$, so that Γ is commensurable with $G(\mathbb{Z})$, has width one at infinity, and contains and normalizes $\Delta = \Gamma_0(nh)$. According to Theorem 6.1.1 the normalized Rademacher sum $T_{\Gamma}^{(-1)}(z)$ is Γ-invariant, so the second condition of Theorem 6.5.1 is satisfied. According to Lemma 6.4.2 there exists a system $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_{\Gamma}\}$ of scaling cosets for Γ such that Γ^{p} contains Δ for each cusp $\mathsf{p} \in \mathsf{P}_{\Gamma}$. Applying Proposition 6.4.3 we see that the normalized Rademacher sum $T_{\Gamma|\mathsf{p}}^{(-1)}(z)$ is Δ-invariant when the scaling coset Σ_{p} of the above scaling coset system is chosen for the definition of $T_{\Gamma|\mathsf{p}}^{(-1)}(z)$. This confirms that the third condition of Theorem 6.5.1 is satisfied. The fourth condition of the main theorem of [CMS04] is that the quotient Γ/Δ have exponent two. Let $\mathsf{r} \in \mathsf{P}_{\Delta}$ be a cusp of Δ that is contained in $\Gamma\cdot\infty$, and let Σ_{r} be the unique right coset of Δ in Γ such that $\mathsf{r} = \Sigma_{\mathsf{r}}\cdot\infty$. Then the exponent two condition implies $\Sigma_{\mathsf{r}}^2 = \Delta$, so that the fourth condition of Theorem 6.5.1 follows from an application of Lemma 6.3.1.

Suppose now that Γ satisfies the four conditions of Theorem 6.5.1. Then Γ satisfies the second condition of [CMS04], and so we have $\Gamma = \Gamma_0(n\|h) + S$ for some $n, h \in \mathbb{Z}^+$, and $S < \mathrm{Ex}(n/h)$, and Γ is a group commensurable with $G(\mathbb{Z})$ that has width one at infinity. Applying Theorem 6.1.1 to the second condition of Theorem 6.5.1 we conclude that Γ has genus zero, and so satisfies the first condition of [CMS04]. Set $\Delta = \Gamma_0(nh)$, so that Γ contains and normalizes Δ. Applying Proposition 6.4.3 to the third condition of Theorem 6.5.1 we see that scaling cosets $\{\Sigma_{\mathsf{p}} \mid \mathsf{p} \in \mathsf{P}_\Gamma\}$ can be chosen for Γ so that Γ^{p} contains Δ for each cusp $\mathsf{p} \in \mathsf{P}_\Gamma$. Applying Lemma 6.4.2 to this fact we conclude that Γ satisfies the fourth condition of the main theorem of [CMS04]. It remains to check that the quotient Γ/Δ has exponent two. Let $\gamma \in \Gamma$ and set $\mathsf{r} = \Delta\gamma \cdot \infty$. If we take $\Sigma_{\mathsf{r}} = \Delta\gamma$ then the fourth condition of Theorem 6.5.1 states that $T^{(-1)}_{\Delta|\mathsf{r}}(z) = T^{(-1)}_{\Delta,\mathsf{r}}(z)$. Proposition 6.3.2 now implies that $\Sigma_{\mathsf{r}}^2 = \Delta$, which in turn implies $\gamma^2 \in \Delta$. This argument applies to arbitrary $\gamma \in \Gamma$ so we conclude that the quotient Γ/Δ has exponent two. This completes the proof. □

Perhaps the most technical condition of Theorem 6.5.1 is the last one. Beyond the monstrous moonshine conjectures, there are the generalized moonshine conjectures of Norton (cf. [Mas87]) which associate genus zero groups to commuting pairs of elements in the Monster. A number of the groups appearing in generalized moonshine do not satisfy the last condition of Theorem 6.5.1, but they all satisfy the second property, and a slight weakening of the first (cf. [Fer93]), and we do not know of any examples that fail to satisfy the third condition of Theorem 6.5.1. It is an interesting question then to determine if the first three conditions of Theorem 6.5.1 furnish a characterization of the groups of generalized moonshine.

7 Gravity

In this section we consider applications of the normalized Rademacher sums to chiral 3d quantum gravity. We should note that the notion of chiral 3d quantum gravity has not yet been defined, so the applications we make, and in particular the conjectures we formulate, are intended to shed light on the very problem of formulating a definition, in addition to elucidating important structural properties that a certain distinguished (and conjectural) example might satisfy. In the absence of a precise definition of 3d quantum gravity some of our discussion in this section is necessarily more speculative.

7.1 First conjecture

It is clear from our results that the Rademacher sums are particularly convenient for understanding the special characteristics of the McKay–Thompson series; in particular, their crucial genus zero property. In order to fully explain the moonshine phenomena, one has to relate these sums to the structure of the vertex operator algebra $V^\natural$. In search for the new relation, one might look again for hints from physics. Very recently, Witten, revisiting 3d quantum gravity, has formulated a number of results and observations [Wit07] including a conjecture about the existence of the 3d quantum gravities with central charges $c_L = c_R$ proportional to

24. In particular, he asserted that the simplest in his list of 3d quantum gravities should be equivalent to the 2d conformal field theory $V^\natural \otimes (V^\natural)^*$. Manschot [Man07] then suggested to consider chiral 3d quantum gravities, the simplest of which, with $c_L = 24$ and $c_R = 0$, should be equivalent exactly to $V^\natural$. Furthermore, Li, Song and Strominger [LSS08] have argued that the chiral gravity possesses stability and consistency, the necessary properties of a sound physical theory. A very recent work [MSS09] by Maloney, Song and Strominger provided further support for the existence of chiral gravities with c_L proportional to 24. These results and observations from physics taken together suggest that there exists an alternative construction of the vertex operator algebra $V^\natural$ that may be viewed as a rigorous version of the simplest chiral 3d quantum gravity with $c = 24$, in the same way as the original construction of $V^\natural$ was interpreted as a rigorous version of the chiral 2d conformal field theory with the partition function $J(z)$. In 3d quantum gravity one expects to obtain the partition function as a sum over minimum points of a 3d quantum gravity action; i.e. over all 3d hyperbolic structures on a solid torus with genus one boundary, whose conformal structure corresponds to the point z on the moduli space. Since all such structures are naturally parameterized by $\Gamma_\infty \backslash PSL_2(\mathbb{Z})$ (cf. [DMMV00]), the partition function should be a kind of Rademacher sum. In fact Manschot and Moore [MM07] argued that the subtraction of constants in (1.2.1) can be explained by a regularization of the partition function of 3d gravity. Our continued Rademacher sums of §4, leading naturally to the modified and normalized Rademacher sums, provide a good candidate for such a regularization.

To obtain the McKay–Thompson series (1.1.3) from the chiral 3d quantum gravity is a more challenging problem even at the heuristic level. However, one may look at these series from a slightly different point of view. It has been shown in [DLM00] that for any $g \in \mathbb{M}$ there exists a unique simple g-twisted $V^\natural$-module $V^\natural_g$, whose partition function is equal to $c(g)T_g(-1/z)$ where $c(g)$ is a constant depending on $g \in \mathbb{M}$. The heuristic analysis of the twisted sector $V^\natural_g$ in [Tui95] strongly supports the general assumption that $c(g) = 1$ for all $g \in \mathbb{M}$. (For more on this assumption please see the discussion of “Hypothesis A_g” in [Car09].)

We now state our first conjecture.

Conjecture 7.1.1. *There exists a family of twisted chiral 3d quantum gravities at central charge $c = 24$ associated with elements of the Monster $g \in \mathbb{M}$ whose partition functions are naturally given by sums over geometries parameterized by $\Gamma_\infty\backslash\Gamma_g$, and these partition functions coincide with the normalized Rademacher sum $T^{(-1)}_\Gamma(-1/z)$, for $\Gamma = \Gamma_g$. Moreover, the untwisted ($g = e$) chiral 3d quantum gravity has a VOA structure isomorphic to $V^\natural$, and the twisted 3d quantum gravity corresponding to $g \in \mathbb{M}$ has a structure of g-twisted $V^\natural$-module isomorphic to that of $V^\natural_g$.*

The special case $g = e$ of Conjecture 7.1.1, corresponding to the untwisted chiral 3d quantum gravity, is strongly supported by the physics literature mentioned above. We will now make a few remarks about the general twisted case. The principle of the 3d quantum gravity/2d CFT correspondence suggests that all the structures in either theory should have appropriate counterparts in the other. The g-twisted sector $V_g^\natural$ is an intrinsic part of the extended theory of the chiral 2d CFT associated to $V^\natural$, and, according to the physical principle, should therefore have analogues in chiral 3d quantum gravity. Basic information about these twisted chiral 3d quantum gravities can be extracted from the Rademacher sums $J_g(z) = T_g(-1/z)$. We'll illustrate the general case with the level two examples, which are the cases that $\Gamma = \Gamma_0(2)$ or $\Gamma_0(2)+$. (In the notation of §2.5 the symbols $\Gamma_0(2)+$ are a shorthand for $\Gamma_0(2) + \mathrm{Ex}(2)$.) As in the case that $g = e$ one expects that each term of the Rademacher sum comes from a classical solution of the corresponding twisted chiral quantum gravity. Every Rademacher sum $J_g(z)$ contains the term $\mathrm{e}(1/z)$, which corresponds to the BTZ black hole solution (cf. [BTZ92]). The other terms arising for $\Gamma = \Gamma_0(2)$ can be characterized among all those arising for $\Gamma = G(\mathbb{Z})$ as the solutions that have the same spin structure as the BTZ black hole on the boundary. The Rademacher sum $J_g(z)$ in this case coincides with the partition function of the Ramond sector in supergravity (cf. [MW07, §7]). The group $\Gamma = \Gamma_0(2)+$ is no longer a subgroup of the modular group $G(\mathbb{Z})$. In this case we have to consider orbifold solutions in addition to smooth solutions; we consider orbifold solutions with a codimension 2 singularity along the defining circle of the solid torus which looks locally like $\mathbb{C}/\langle\, \mathrm{e}(1/2)\rangle$ (cf. [MW07, §2]), and we still impose the same spin structure boundary condition. Note that on any space X, choices of spin structure on X (when they exist) are in bijective correspondence with a particular family of double covers of X (cf. e.g. [LM89]). Such a double cover smooths the $\mathbb{Z}/2$-orbifold singularity. One may view the spin structure as a $\mathbb{Z}/2$-structure, and for a general level N group Γ, spin structures are replaced with $\mathbb{Z}/N$-structures. One can also allow singularities along the defining circle of the solid torus that look locally like $\mathbb{C}/\langle\, \mathrm{e}(1/N)\rangle$. In general, a choice of twisted gravity with $\mathbb{Z}/N$-structure group should impose an $n\|h + S$ equivalence for solid tori, for some n and h with $N = nh$ and h a divisor of 24 (cf. Theorem 6.2.3), which yields the Rademacher sum $J_g(z)$, for a corresponding $g \in \mathbb{M}$, as the saddle point approximation of the twisted quantum gravity partition function.

We expect that the conjectural twisted chiral 3d quantum gravity construction of $V_g^\natural$ will imply the Γ_g-invariance of its partition function, in analogy with the way in which 2d CFTs are found to have modular invariant partition functions (cf. [Wit87]). Then our conjecture, in combination with our results on Rademacher sums, will naturally imply the genus zero conjecture of Conway–Norton (cf. [CN79]), and most importantly, will reveal the geometric nature of moonshine. In particular, the solution to Conjecture 7.1.1 will answer the genus zero question: why the discrete groups attached to the Monster via monstrous moonshine all have genus zero.

We also expect that the conditions on Rademacher sums in our reformulation Theorem 6.5.1, of the group theoretic characterization of the groups of monstrous moonshine due to Conway–McKay–Sebbar, will also find an interpretation as consistency conditions for twisted chiral 3d quantum gravities. In particular, the completeness of the family of twisted chiral 3d quantum gravities at central charge 24 associated with elements of the Monster group $\mathbb{M}$ should have a deep meaning in quantum gravity.

Our Conjecture 7.1.1 should have significance for the future development of analytic number theory also. We already know from 2d CFT that various constructions yield remarkable number theoretic identities. Our conjecture implies that the development of the theory of 3d quantum gravity might encompass a whole new family of number theoretic results such as the theory of Rademacher sums studied in this work. In §7.3 we'll show that the Hecke operators also admit an interpretation in terms of quantum gravity. In preparation for this we will recall in the next section some results about a certain class of generalized Kac–Moody algebras and their representations.

7.2 Monstrous Lie algebras

In his paper [Bor92], Borcherds constructed a remarkable generalized Kac–Moody algebra (*GKM algebra*) $\mathfrak{m}$, called the *Monster Lie algebra*, which plays a key rôle in his approach to monstrous moonshine. He also defined GKM superalgebras $\mathfrak{m}'_g$, for each $g \in \mathbb{M}$, by using the McKay–Thompson series $T_g(z)$ to specify the simple roots. In this section we'll consider another family of GKM algebras $\mathfrak{m}_g$, parameterized by elements $g \in \mathbb{M}$, with simple roots specified by the functions $J_g(z) = T_g(-1/z)$. Since the Fourier coefficients of the $J_g(z)$ are non-negative integers, the algebras $\mathfrak{m}_g$ are purely even. We will also study Verma modules for the GKM algebras $\mathfrak{m}_g$.

In the following section §7.3, we will explain how these algebraic structures arise from the Rademacher sums, and the conjectural twisted chiral 3d quantum gravities. This will lead us to a further extension of our first conjecture.

The algebras $\mathfrak{m}_g$ were introduced and studied by Carnahan in [Car07] and were discovered independently by the second author. We will now recall the relevant results of [Car07].

Let $L = II_{1,1}$ denote a copy of the unique even self-dual Lorentzian lattice of rank 2. We identify L with the group $\mathbb{Z} \times \mathbb{Z}$ and set the norm of the pair (m, n) to be $-2mn$. For any positive integer N we denote the sublattice $\mathbb{Z} \times N\mathbb{Z}$ by $L(N)$, and we let $L(N)^\vee$ denote the dual lattice $\frac{1}{N}\mathbb{Z} \times \mathbb{Z}$. Let V_L denote the vertex operator algebra (VOA) associated to the lattice L (cf. [FLM88]),

$$V_L = \bigoplus_{(m,n)\in L} V_L^{(m,n)}, \tag{7.2.1}$$

and let h_N denote the automorphism of V_L which acts as multiplication by $\mathrm{e}(n/N)$ on the subspace $V_L^{(m,n)}$. Then $V_{L(N)}$ is a vertex operator subalgebra of $V_{L(N)}$ fixed by h_N. Following [DL93] each coset $L + (k/N, 0)$ of L in $L(N)^\vee$ defines an h_N^k-twisted module for V_L. We denote this twisted module by $V_{L+(k/N,0)}$, and regard it as graded in the natural way by $L + (k/N, 0)$.

For $g \in \mathbb{M}$ such that the level of Γ_g is N (i.e. $N = nh$ in the notation of [CN79]) define $W_g^\natural$

to be the following $L(N)^\vee$-graded space invariant under $\langle g\rangle$.

$$W_g^\natural = \bigoplus_{k=0}^{N-1} \left(V_{g^k}^\natural \otimes V_{L+(k/N,0)} \right)^{\langle g\rangle} \tag{7.2.2}$$

The summand corresponding to $k=0$ has a natural VOA structure, while the other summands are naturally modules for this VOA. There also exists a unique (up to scalar) intertwining operator between the product of the k-th and l-th summands and the $(k+l)$-th summand. Carnahan asserts that there is a consistent choice of these constants such that one has

Theorem 7.2.1. *The space $W_g^\natural$ naturally admits a VOA structure.*

By construction $W_g^\natural$ has rank 26 and one can define the semi-infinite cohomology of the Virasoro algebra with coefficients in $W_g^\natural$. The the first semi-infinite cohomology group acquires a Lie algebra structure according to [LZ95]. In this way we obtain a *monstrous Lie algebra* $\mathfrak{m}_g$ for each $g \in \mathbb{M}$.

$$\mathfrak{m}_g = H^{\infty/2+1}(W_g^\natural) \tag{7.2.3}$$

An alternative construction of $\mathfrak{m}_g$ may be based on the no-ghost theorem of the 26-dimensional bosonic string (cf. [Fre85], [Bor92]). For $g=e$, the Lie algebra $\mathfrak{m}_e$ is the original Monster Lie algebra of Borcherds [Bor92].

The Lie algebra $\mathfrak{m}_g$ inherits an $L(N)^\vee$-grading

$$\mathfrak{m}_g = \bigoplus_{m,n\in\mathbb{Z}} \mathfrak{m}_g^{(m/N,n)}. \tag{7.2.4}$$

The same argument as in the case $g=e$ applied to the semi-infinite cohomology or no-ghost theorem construction of $\mathfrak{m}_g$ yields natural isomorphisms

$$\mathfrak{m}_g^{(m/N,n)} \cong V_{g^{-m},n/N}^\natural(mn/N) \tag{7.2.5}$$

where $V_{g^{-m},n/N}^\natural$ is the subspace of $V_{g^{-m}}^\natural$ upon which g acts by $\mathrm{e}(n/N)$ and $V_{g^{-m},n/N}^\natural(mn/N)$ is the L_0-eigenspace of $V_{g^{-m},n/N}^\natural$ with eigenvalue $mn/N+1$.

Remark 7.2.2. Let us write V^Λ for the vertex operator algebra associated to the Leech lattice. In his paper [Tui95] Tuite gave strong evidence that the genus zero property of monstrous moonshine is equivalent to the following duality isomorphisms,

$$V_{g^m,n/N}^\natural \cong V_{g^n,m/N}^\natural, \tag{7.2.6}$$

$$V_{g^m,n/N}^\natural \cong V_{h^n,m/N}^\Lambda, \tag{7.2.7}$$

the first holding in the case that g is Fricke, and the second in the case that g is non-Fricke. In the second case we may take g to lie in a subgroup $2^{1+24}.\,Co_1$ of $\mathbb{M}$ and the h in (7.2.7) denotes an element of $2^{24}.2.\,Co_1$, a subgroup of $\mathrm{Aut}(V^\Lambda)$, such that h and g project to the same conjugacy class of Co_1. Thus the Tuite duality in the Fricke case implies the symmetry

$$\mathfrak{m}_g^{(m/N,n)} \cong \mathfrak{m}_g^{(n/N,m)} \tag{7.2.8}$$

of the monstrous Lie algebra $\mathfrak{m}_g$, as is known to hold in the case $g = e$ studied in [Bor92]. In the non-Fricke case the Tuite duality suggests an alternative construction of the VOA $W_g^\natural$ as

$$W_g^\natural = \bigoplus_{k=0}^{N-1} \left(V_{h^k}^\Lambda \otimes V_{\tilde{L}+(0,k)} \right)^{\langle h \rangle} \tag{7.2.9}$$

where $\tilde{L} = \frac{1}{N}\mathbb{Z} \times N\mathbb{Z}$.

We now re-scale the grading in $W_g^\natural$ and $\mathfrak{m}_g$ by interchanging $(m/N, n)$ with $(m, n/N)$, and we set

$$\mathfrak{m}_g^m = \bigoplus_{n \in \frac{1}{N}\mathbb{Z}} \mathfrak{m}_g^{(m,n/N)} \tag{7.2.10}$$

for $m \in \mathbb{Z}$. Then (7.2.5) implies isomorphisms of graded spaces

$$\mathfrak{m}_g^1 \cong V_{g^{-1}}^\natural, \quad \mathfrak{m}_g^{-1} \cong V_g^\natural, \tag{7.2.11}$$

and we also have

$$\mathfrak{m}_g^0 = \mathfrak{m}_g^{(0,0)} = \mathbb{C}c \oplus \mathbb{C}d, \tag{7.2.12}$$

where c and d are the degree operators for the re-scaled grading.

In [Car07] Carnahan shows that $\mathfrak{m}_g$ is a GKM algebra, thus generalizing a result of [Bor92]. When g is Fricke, $\mathfrak{m}_g$ has a structure very similar to that of the original Borcherds Monster Lie algebra $\mathfrak{m} = \mathfrak{m}_e$. In particular, it has one real simple root and all simple roots correspond to a basis for $V_g^\natural$. This implies that $\mathfrak{m}_g$ can be reconstructed from the subalgebra

$$\mathfrak{m}_g^1 \oplus \mathfrak{m}_g^0 \oplus \mathfrak{m}_g^{-1} \tag{7.2.13}$$

which is called the *local subalgebra of* $\mathfrak{m}_g$ (cf. [FF83]).

Remark 7.2.3. When g is non-Fricke $\mathfrak{m}_g$ does not have real simple roots, and in general it cannot be reconstructed from its local subalgebra (7.2.13). However the Tuite duality suggests that if one considers the second grading of $\mathfrak{m}_g$ then the local subalgebra consists of V_h^Λ and $V_{h^{-1}}^\Lambda$ and the h-twisted Heisenberg algebra. This local subalgebra is expected to generate $\mathfrak{m}_g$ in the case that g is non-Fricke.

In his thesis [Car07] Carnahan also obtained remarkable generalizations of the Borcherds identities for each of the monstrous Lie algebras $\mathfrak{m}_g$; viz.,

$$p(T_g(w) - J_g(z)) = \prod_{m\in\mathbb{Z}^+,n\in\mathbb{Z}} (1-p^m q^{n/N})^{c_{g^m,n/N}(mn/N)} \tag{7.2.14}$$

where we have set

$$c_{g^m,n/N}(mn/N) = \dim V^\natural_{g^m,n/N}(mn/N). \tag{7.2.15}$$

When g is Fricke the invariance of $T_g(w)$ under the Fricke involution and a re-scaling $w \mapsto Nw$ yields

$$p(J_g(Nw) - J_g(Nz)) = \prod_{m\in\mathbb{Z}^+,n\in\mathbb{Z}} (1-p^m q^n)^{c_{g^m,n/N}(mn/N)}. \tag{7.2.16}$$

Dividing both sides of (7.2.16) by $(1-pq^{-1})$ we obtain expressions which are invariant under the transposition of p with q. This entails the identity

$$c_{g^m,n/N}(mn/N) = c_{g^m,m/N}(mn/N) \tag{7.2.17}$$

which is in agreement with the Tuite duality of (7.2.6). The identity (7.2.14) implies an unexpected formula for the graded dimension of the Verma module with trivial character $\mathcal{V}_g$ associated to g.

$$\mathcal{V}_g = \mathcal{U}(\mathfrak{m}_g^-) \tag{7.2.18}$$

In fact the standard product formula

$$\operatorname{gdim} \mathcal{V}_g = \prod_{m\in\mathbb{Z}^+,n\in\mathbb{Z}} (1-p^m q^{n/N})^{-c_{g^m,n/N}(mn/N)} \tag{7.2.19}$$

for the graded dimension of $\mathcal{V}_g$ follows from (7.2.5) and (7.2.15). The identity (7.2.14) yields the alternative expression

$$\operatorname{gdim} \mathcal{V}_g = \frac{1}{p(T_g(w) - J_g(z))}. \tag{7.2.20}$$

This expression allows us to view the bi-graded dimension as a meromorphic function on $\mathbb{H}\times\mathbb{H}$.

Thanks to the properties of principal moduli for curves of genus zero, the expression (7.2.20) is singular at the point (w,z) precisely when $\Gamma_g \cdot w = \Gamma_g \cdot (-1/z)$. In §7.3 we will give an interpretation of this fact in the setting of three dimensional quantum gravity.

With the goal of making such an interpretation in mind we propose to rewrite the first expression (7.2.19) for the bi-graded dimension of the Verma module $\mathcal{V}_g$ as

$$\operatorname{gdim} \mathcal{V}_g = \exp\left(\sum_{m\in\mathbb{Z}^+} \hat{T}_g(m) J_g(z) p^m \right) \tag{7.2.21}$$

using generalized Hecke operators $\hat{T}_g(m)$ such as those discussed in §5.5; we have verified there the existence of such operators for $g = e$ and we conjecture that suitable $\hat{T}_g(m)$ exist for every $g \in \mathbb{M}$. Comparing then with the generating functions

$$Z_g(p,q) = \sum_{m\in\mathbb{Z}^+} m\hat{T}_g(m) J_g(z) p^m \tag{7.2.22}$$

of the higher order Rademacher sums $J_g^{(-m)}(z)$ we obtain the following expression.

$$Z_g(p,q) = p\partial_p \log(\operatorname{gdim} \mathcal{V}_g) \tag{7.2.23}$$

An identity of this form is well-known in the theory of symmetric functions (cf. [Mac95]), where it serves to relate the generating function of the power symmetric functions

$$p_n = \sum_{1\leq i} x_i^n \tag{7.2.24}$$

with that of the complete symmetric functions

$$h_n = \sum_{1\leq i_1 \leq \cdots \leq i_n} \prod_{k=1}^{n} x_{i_k}. \tag{7.2.25}$$

Thus the graded dimensions of the Verma modules $\mathcal{V}_g$ may be viewed as *complete Rademacher sums*, where the rôle of the variables x_i, for $i \in \mathbb{Z}^+$, in (7.2.24) and (7.2.25) is taken up by the exponential expressions $\mathrm{e}(\llbracket\gamma \cdot z')$ for $z' = -1/z$ and $\llbracket\gamma \in \llbracket\Gamma_g$.

Since all the coefficients appearing in the bi-graded dimension $Z_g(p,q)$ are non-negative integers, one may expect to naturally find a bi-graded vector space, and possibly even an $\mathfrak{m}_g$-module, with bi-graded dimension given by $Z_g(p,q)$. The next result suggests that one can expect to find such a space within $\mathcal{V}_g$.

Given two elements $F = \sum F_{m,n} p^m q^n$ and $G = \sum G_{m,n} p^m q^n$ in $\mathbb{Z}(\!(q)\!)[[p]]$, let us write $F \leq G$ in the case that $F_{m,n} \leq G_{m,n}$ for all $m,n \in \mathbb{Z}$.

Proposition 7.2.4. *Let $g \in \mathbb{M}$. If g is Fricke then we have $Z_g(p,q) \leq \operatorname{gdim} \mathcal{V}_g$.*

Proof. Dividing both sides of (7.2.16) by $(1-pq^{-1})$ we obtain

$$1 - \sum_{m,n\in\mathbb{Z}^+} c_{g,\frac{m+n-1}{N}}\left(\frac{m+n-1}{N}\right)p^m q^n = \prod_{m,n\in\mathbb{Z}^+}(1-p^m q^n)^{c_{g^m,n/N}(mn/N)}. \tag{7.2.26}$$

We denote the sum in the left hand side by Σ^+ and the product in the right hand side by Π^+. Then we have

$$Z_g(p,q) = -p\partial_p \log(1-pq^{-1})\Pi^+ = \frac{pq^{-1}}{1-pq^{-1}} + \frac{p\partial_p\Sigma^+}{1-\Sigma^+}. \tag{7.2.27}$$

On the other hand

$$\operatorname{gdim}\mathcal{V}_g = \frac{1}{1-pq^{-1}}\frac{1}{1-\Sigma^+} = \frac{1}{1-pq^{-1}} + \frac{\Sigma^+}{(1-pq^{-1})(1-\Sigma^+)}. \tag{7.2.28}$$

Thus it is sufficient to show that

$$p\partial_p(\Sigma^+) \leq \frac{1}{1-pq^{-1}}\Sigma^+, \tag{7.2.29}$$

but this follows from the observation that both sides of (7.2.29) have nonnegative coefficients and the left hand side may be obtained from the right hand side by deleting the terms which are singular or constant with respect to q. □

To formulate a conjecture on the nature of the subspaces of $\mathcal{V}_g$ with bi-graded dimensions $Z_g(p,q)$, as well as the origin of the higher order Rademacher sums, we turn again, in the next section, to the structures of twisted chiral 3d quantum gravity.

7.3 Second conjecture

We have shown in §5.5 that the higher order Rademacher sums recover the action of the Hecke operators on the first order Rademacher sums. In view of the conjectural relation between first order Rademacher sums and chiral 3d quantum gravities at central charge $c = 24$, we may consider an analogous interpretation for the higher order Rademacher sums given by $m(\hat{T}_g(m)J_g)(z)$ for $g \in \mathbb{M}$ and $m \in \mathbb{Z}^+$. Since all the Fourier coefficients of these expressions are non-negative integers, one might guess that there exist families of twisted chiral 3d quantum gravities for all central charges $c = 24m$, for $m \in \mathbb{Z}^+$. In particular, the untwisted ($g = e$) chiral 3d quantum gravities possess the structure of certain extremal vertex algebras with the Monster symmetry.

The aforementioned class of extremal vertex algebras, with the addition of a Virasoro element, has been originally conjectured by Witten [Wit07]. However, it was shown in [Gai08] and

[Hoe08] that the addition of a Virasoro element precludes a non-trivial action of the Monster group. The higher order Rademacher sums also point to the partition functions without elements of spin 2. We may modify Witten's conjecture by asking for the existence of extremal vertex algebras with partition functions given by the function $J^{(-m)}(z)$ for $m > 0$, related to the higher order Rademacher sums via $J^{(-m)}(z) = Q_{\Gamma}^{(-m)}(z) - \mathrm{c}_{\Gamma}(-m, 0)$, with $\Gamma = PSL_2(\mathbb{Z})$. The validity of this conjecture remains open.

Note that the Virasoro algebra, even when not represented by any actual state, may still act on a vertex algebra with partition function $J^{(-m)}(z)$. Regardless of whether or not spaces $V^{(-m)}$ with partition function $J^{(-m)}(z)$ admit natural vertex algebra structures, they certainly naturally inherit actions of the Monster group, since the actions of Hecke operators $m\hat{T}(m)$ on the Fourier coefficients of $J(z)$ may be interpreted as actions on representations of $\mathbb{M}$ (cf. [JLW95]).

It is natural to enquire as to the meaning of the spaces $V_g^{(-m)}$, for $m \in \mathbb{Z}^+$ and $g \in \mathbb{M}$, with graded dimension $J_g^{(-m)}(z)$. In light of our First Conjecture, it is unlikely that the spaces $V^{(-m)}$ can represent states of a three dimensional quantum gravity with $c = 24m$, since a Virasoro element is generally not present. Thus we have to conclude that the three dimensional quantum gravity $V^{(-1)} = V^{\natural}$ is the only viable candidate. Then one might view the spaces $V^{(-m)}$ for $m > 1$ as some higher overtones of the basic $m = 1$ theory.

In fact, we have shown in §5.5 that the action of the classical Hecke operator $\hat{T}(m)$ on the Rademacher sum (1.2.1) yields a sum over $\Gamma_\infty \backslash M(m)$, where $M(m)$ denotes the (image in $G(\mathbb{Q})$ of the) set of 2×2 matrices with integral entries and determinant m. This sum may be viewed as the *m-instanton correction* of the partition function in three dimensional quantum gravity. By *n-instanton* in this context one understands an elliptic curve which admits a holomorphic map of degree n into a given elliptic curve $E_{[\![z}$ (cf. [DMVV97]). Thus the sum over $\Gamma_\infty \backslash M(n)$ becomes a sum over all three dimensional hyperbolic structures on a solid torus with genus one boundary whose conformal structure corresponds to an n-instanton on $E_{[\![z}$. The same phenomena is expected for all the twisted 3d quantum gravities corresponding to elements $g \in \mathbb{M}$.

Collecting all the m-instanton contributions in one generating function we then obtain a partition function

$$Z_g(p,q) = \sum_{m \in \mathbb{Z}^+} m(\hat{T}_g(m)J_g)(z)p^m \tag{7.3.1}$$

depending on two variables, for each $g \in \mathbb{M}$. This partition function $Z_g(p,q)$ may be viewed as a part of the full partition function

$$\tilde{Z}_g(p,q) = \exp\left(\sum_{m \in \mathbb{Z}^+} (\hat{T}_g(m)J_g)(z)p^m\right) \tag{7.3.2}$$

of the stringy second quantization of $V_g^\natural$, introduced in [DMVV97] in the untwisted case $g = e$. (See also [Tui08].) We will now generalize their construction to an arbitrary twisted module $V_g^\natural$.

Let V be an arbitrary vertex operator algebra; in our case $V = V^\natural$. For n a positive integer, let $V^{\otimes n}$ denote the tensor product of n copies of V, and let Z_n be the group of cyclic permutations of the factors generated by an n-cycle $\sigma_n = (1, 2, \cdots, n)$. Let g be an automorphism of V of order N, then $g \times \sigma_n$ is an automorphism of $V^{\otimes n}$ of order nN/d where $d = (n, N)$. To the pair $(V^{\otimes n}, g \times \sigma_n)$ is canonically associated a twisted module (cf. [BDM02]), which we denote $V_{(g,n)}$. The group $Z_{g,n} = \langle g \times \sigma_n \rangle$ acts naturally on $V_{(g,n)}$. We write $V_{(g,n)}^{Z_{g,n}}$ for the $Z_{g,n}$-invariant subspace. The *stringy second quantization of the twisted module* V_g, to be denoted $\mathbb{S}\, V_g$, is, by definition, the space

$$\mathbb{S}\, V_g = \bigoplus_{\lambda \in \mathcal{P}} \bigotimes_{r>0} S^{m_r} V_{(g,r)}^{Z_{g,r}}, \tag{7.3.3}$$

where the sum is taken over all partitions λ with m_r parts of length $r > 0$. The space $\mathbb{S}\, V_g$ is doubly graded: by the degrees of products of twisted sectors and by the value of $|\lambda|$. For $g = e$ our definition of the stringy second quantization coincides with that of [DMVV97].

Next we establish a relation between the second quantization $\mathbb{S}\, V_g^\natural$ and the Verma module $\mathcal{V}_g$. First note the canonical isomorphism

$$\mathcal{U}(\mathfrak{m}_g^-) \cong S(\mathfrak{m}_g^-). \tag{7.3.4}$$

Then we obtain the following result.

Theorem 7.3.1. *There is a canonical isomorphism of bi-graded vector spaces*

$$\mathbb{S}\, V_g^\natural \cong S(\mathfrak{m}_g^-). \tag{7.3.5}$$

Proof. The isomorphism (7.3.5) follows from the existence of isomorphisms

$$(V_{(g,n)}^\natural)^{Z_{g,n}} \cong \mathfrak{m}_g^{-n} \tag{7.3.6}$$

of bi-graded vector spaces for each $n \in \mathbb{Z}^+$. To establish the existence of these we will use the isomorphism

$$V_{(g,n)}^\natural \cong V_{g^n}^\natural \tag{7.3.7}$$

of [BDM02].

We wish to show that the subspace of $V^\natural_{g^n}$ fixed by the operator

$$g \times \mathrm{e}((L_0 - 1)/n) \tag{7.3.8}$$

is naturally isomorphic to $\mathfrak{m}_g^n$. Consider the invariant subspace of $V^\natural_{g^n}$ with respect to the N-th power of the operator (7.3.8), which may be expressed as $\mathrm{e}((L_0 - 1)N/n)$. Its action on the graded subspace $V^\natural_{g^n}(kd/N)$ for $k \in \mathbb{Z}$ is multiplication by $\mathrm{e}(kd/n)$, and this scalar is 1 if and only if $k = (n/d)m$ for some $m \in \mathbb{Z}$. Thus we are only concerned with the graded subspaces $V^\natural_{g^n}(nm/N)$ where $m \in \mathbb{Z}$. To find the invariant subspaces with respect to the operator (7.3.8) we consider its action on the subspaces of the form $V^\natural_{g^n,m'/N}(mn/N)$. The action is scalar multiplication by $\mathrm{e}(-m'/N)\,\mathrm{e}(m/N)$ and is therefore trivial if and only if $m \equiv m' \pmod{N}$. Thus the full subspace of $V^\natural_{g^n}$ invariant under (7.3.8) is given by

$$\bigoplus_{m\in\mathbb{Z}} V^\natural_{g^n,m/N}(mn/N) \cong \mathfrak{m}_g^{-n}. \tag{7.3.9}$$

To complete the proof we note that the group generated by the action of the operator (7.3.8) has order nN/d, and under the isomorphism (7.3.7) its action recovers that of $Z_{g,n}$. □

The isomorphism of Theorem 7.3.1 suggests that the stringy second quantization $\mathbb{S}\,V^\natural_g$ admits an action by the Lie algebra $\mathfrak{m}_g$ and thus provides a Fock space realization of the Verma module $\mathcal{V}_g$. When $\mathfrak{m}_g$ is generated by its local subalgebra (7.2.13), as happens in the Fricke case, it is sufficient to describe the action of brackets $[\mathfrak{m}_g^{\pm 1}, \mathfrak{m}_g^{-n}] \subset \mathfrak{m}_g^{-n\pm 1}$ via the action of $V^\natural_{g^{\pm 1}}$ on the twisted sector for $((V^\natural)^{\otimes n}, g \times \sigma_n)$. Note also that the removal of the subspaces corresponding to the terms which are singular or constant with respect to q, as in the proof of Proposition 7.2.4, will yield an embedding of the instanton subspace $\mathcal{I}_g$ inside the stringy second quantization of $V^\natural_g$.

Our interpretation of the higher order Rademacher sums and the complete Rademacher sums, and their relation to the second quantization of the moonshine vertex operator algebra and its twisted modules, as well as the monstrous Lie algebras and their Verma modules, suggests the following extension of our first conjecture, Conjecture 7.1.1.

Conjecture 7.3.2. *Assuming the existence of a family of g-twisted chiral 3d quantum gravities for $g \in \mathbb{M}$, having the properties stated in Conjecture 7.1.1, there also exists a natural geometric interpretation of the family of twisted monstrous Lie algebras $\mathfrak{m}_g$, the instanton spaces $\mathcal{I}_g$, and the denominator formulas for all $g \in \mathbb{M}$ via the second quantization of the corresponding g-twisted chiral three 3d quantum gravities.*

First of all, the partition function of the second quantized g-twisted chiral 3d quantum gravity should depend on two modular parameters w and z, and should respect the symmetry that interchanges w with $-1/z$. Second, there should be a quantum gravity theoretic explanation for why these partition functions are singular precisely when w and $-1/z$ belong to the same orbit of Γ_g. This will explain the remarkable denominator formulas (7.2.14) of Carnahan, and ultimately, the fundamental rôle of the principal moduli in the moonshine Conjectures.

Where can we find an appropriate setting for all the structures that appear in our Second Conjecture? The first answer that might come to mind is the 26 dimensional bosonic string theory. It was known for a long time that the physical space has a Lie algebra structure (cf. [Fre85], [GO85]), which can also be recovered from the semi-infinite cohomology (cf. [FGZ86], [LZ95]). However, in this way one can only get a fake Monster Lie algebra (cf. [Bor92]) and there is no (straight forward) geometric way to turn it into the real Monster Lie algebra that appears in the three dimensional quantum gravity approach. Also, the 26 dimensions of the bosonic string have nothing to do with the 3 dimensions of the quantum gravity. It is still possible that some constructions of string theory can be applied in the 3d quantum gravity setting. In particular, one can expect to identify the negative part $\mathfrak{m}^-$ of the Monster Lie algebra in the space of the second quantization of $V^\natural$ with a certain version of BPS states (cf. [HM96]).

Clearly, our results on Rademacher sums, their relation to the moonshine module, monstrous Lie algebras, and 3d quantum gravity, admit a super-counterpart [DF09]. In this case the promise of relationships with the structures of the 10 dimensional superstring is even more tempting, since the latter has played such a prominent rôle in physics over the past 25 years. However, in spite of the of the remarkable mathematical similarity, the 10 of the ten dimensional superstring and the 3 of the 3d quantum supergravity emerging from the super-counterpart of the Rademacher sums have different geometric meaning and are not related by a compactification of any kind.

The most fascinating fact about the 3d quantum supergravity is that while it is similar, but not directly related, to the initial 3d quantum gravity, the corresponding second quantized theories are directly related. The higher order Rademacher sums point to a larger second quantized space in the super-case; a space which contains the second quantized space we considered above.

Then what can these second quantized 3d quantum gravities mean in physics? Naïvely, any quantization can be interpreted as a categorification, which lifts a given theory one dimension up. One can then wonder if the second quantization in our case might point to an extreme sector of certain four dimensional quantum gravities, such as the extreme Kerr black hole, which was recently found in [GHSS08] also to be dual to a chiral 2d CFT. In this case the Monster (or moonshine) might be the answer to the perpetual question of what is behind the letter M in the theory that has not yet revealed its true name.

Acknowledgement

We are very grateful to the referees for many helpful comments and suggestions.

References

[AS64] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[BDM02] Katrina Barron, Chongying Dong, and Geoffrey Mason. Twisted sectors for tensor product vertex operator algebras associated to permutation groups. *Comm. Math. Phys.*, 227(2):349–384, 2002.

[Bor92] Richard E. Borcherds. Monstrous moonshine and monstrous Lie superalgebras. *Invent. Math.*, 109(2):405–444, 1992.

[BTZ92] Máximo Bañados, Claudio Teitelboim, and Jorge Zanelli. Black hole in three-dimensional spacetime. *Phys. Rev. Lett.*, 69(13):1849–1851, 1992.

[Car07] Scott Carnahan. *Monstrous Lie algebras and generalized moonshine*. PhD thesis, University of California, Berkeley, 2007.

[Car09] Scott Carnahan. Generalized Moonshine II: Borcherds products. arXiv:0908.4223.

[CMS04] John Conway, John McKay, and Abdellah Sebbar. On the discrete groups of Moonshine. *Proc. Amer. Math. Soc.*, 132:2233–2240, 2004.

[CN79] J. H. Conway and S. P. Norton. Monstrous moonshine. *Bull. London Math. Soc.*, 11(3):308–339, 1979.

[DF09] John F. Duncan and Igor B. Frenkel. Rademacher sums and super-moonshine. In preparation.

[DL93] Chongying Dong and James Lepowsky. *Generalized vertex algebras and relative vertex operators*, volume 112 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1993.

[DLM00] Chongying Dong, Haisheng Li, and Geoffrey Mason. Modular-invariance of trace functions in orbifold theory and generalized Moonshine. *Communications in Mathematical Physics*, 214:1–56, 2000.

[DMMV00] Robbert Dijkgraaf, Juan Maldacena, Gregory Moore, and Erik Verlinde. A black hole Farey tail. arXiv:hep-th/0005003v3, April 2000.

[DMVV97] Robbert Dijkgraaf, Gregory Moore, Erik Verlinde, and Herman Verlinde. Elliptic genera of symmetric products and second quantized strings. *Comm. Math. Phys.*, 185(1):197–209, 1997.

[Fer93] Charles R. Ferenbaugh. The genus-zero problem for $n|h$-type groups. *Duke Math. J.*, 72(1):31–63, 1993.

[FF83] Alex J. Feingold and Igor B. Frenkel. A hyperbolic Kac-Moody algebra and the theory of Siegel modular forms of genus 2. *Math. Ann.*, 263(1):87–144, 1983.

[FGZ86] I. B. Frenkel, H. Garland, and G. J. Zuckerman. Semi-infinite cohomology and string theory. *Proc. Nat. Acad. Sci. U.S.A.*, 83(22):8442–8446, 1986.

[FLM88] Igor Frenkel, James Lepowsky, and Arne Meurman. *Vertex operator algebras and the Monster*, volume 134 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1988.

[Fre85] I. B. Frenkel. Representations of Kac-Moody algebras and dual resonance models. In *Applications of group theory in physics and mathematical physics (Chicago, 1982)*, volume 21 of *Lectures in Appl. Math.*, pages 325–353. Amer. Math. Soc., Providence, RI, 1985.

[Gai08] Davide Gaiotto. Monster symmetry and extremal CFTs. arXiv:0801.0988v1, January 2008.

[GHSS08] Monica Guica, Thomas Hartman, Wei Song, and Andrew Strominger. The Kerr/CFT correspondence. arXiv:0809.4266v1, September 2008.

[GO85] P. Goddard and D. Olive. Algebras, lattices and strings. In *Vertex operators in mathematics and physics (Berkeley, Calif., 1983)*, volume 3 of *Math. Sci. Res. Inst. Publ.*, pages 51–96. Springer, New York, 1985.

[GS83] D. Goldfeld and P. Sarnak. Sums of Kloosterman sums. *Invent. Math.*, 71(2):243–250, 1983.

[Hej83] Dennis A. Hejhal. *The Selberg trace formula for* PSL(2, $\mathbf{R}$). *Vol. 2*, volume 1001 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.

[HM96] Jeffrey A. Harvey and Gregory Moore. Algebras, BPS states, and strings. *Nuclear Phys. B*, 463(2-3):315–368, 1996.

[Hoe08] G. Hoehn. Self-dual vertex operator superalgebras of large minimal weight. arXiv:0801.1822v1, January 2008.

[HR64] G. H. Hardy and M. Riesz. *The general theory of Dirichlet's series.* Cambridge Tracts in Mathematics and Mathematical Physics, No. 18. Stechert-Hafner, Inc., New York, 1964.

[Iwa97] Henryk Iwaniec. *Topics in classical automorphic forms*, volume 17 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1997.

[Iwa02] Henryk Iwaniec. *Spectral methods of automorphic forms*, volume 53 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2002.

[JLW95] E. Jurisich, J. Lepowsky, and R. L. Wilson. Realizations of the Monster Lie algebra. *Selecta Math. (N.S.)*, 1(1):129–161, 1995.

[Kno61a] Marvin Isadore Knopp. Construction of a class of modular functions and forms. *Pacific J. Math.*, 11:275–293, 1961.

[Kno61b] Marvin Isadore Knopp. Construction of a class of modular functions and forms. II. *Pacific J. Math.*, 11:661–678, 1961.

[Kno62a] Marvin Isadore Knopp. Construction of automorphic forms on H-groups and supplementary Fourier series. *Trans. Amer. Math. Soc.*, 103:168–188, 1962.

[Kno62b] Marvin Isadore Knopp. On abelian integrals of the second kind and modular functions. *Amer. J. Math.*, 84:615–628, 1962.

[Kno86] M. I. Knopp. On the Fourier coefficients of small positive powers of $\theta(\tau)$. *Invent. Math.*, 85(1):165–183, 1986.

[Kno89] Marvin I. Knopp. On the Fourier coefficients of cusp forms having small positive weight. In *Theta functions—Bowdoin 1987, Part 2 (Brunswick, ME, 1987)*, volume 49 of *Proc. Sympos. Pure Math.*, pages 111–127. Amer. Math. Soc., Providence, RI, 1989.

[Kno90] Marvin I. Knopp. Rademacher on $J(\tau)$, Poincaré series of nonpositive weights and the Eichler cohomology. *Notices Amer. Math. Soc.*, 37(4):385–393, 1990.

[KR01] Marvin Knopp and Sinai Robins. Easy proofs of Riemann's functional equation for $\zeta(s)$ and of Lipschitz summation. *Proc. Amer. Math. Soc.*, 129(7):1915–1922 (electronic), 2001.

[Kub73] Tomio Kubota. *Elementary theory of Eisenstein series.* Kodansha Ltd., Tokyo, 1973.

[LM89] H. Blaine Lawson, Jr. and Marie-Louise Michelsohn. *Spin geometry*, volume 38 of *Princeton Mathematical Series.* Princeton University Press, Princeton, NJ, 1989.

[LSS08] Wei Li, Wei Song, and Andrew Strominger. Chiral gravity in three dimensions. *J. High Energy Phys.*, (4):082, 15, 2008.

[LZ95] Bong H. Lian and Gregg J. Zuckerman. Moonshine Cohomology. arXiv:q-alg/9501015v1, January 1995.

[Mac95] I. G. MacDonald. *Symmetric functions and Hall polynomials.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[Man07] Jan Manschot. AdS_3 partition functions reconstructed. *J. High Energy Phys.*, (10):103, 7, 2007.

[Mas87] Geoffrey Mason. Finite groups and modular functions. In *The Arcata Conference on Representations of Finite Groups (Arcata, Calif., 1986)*, volume 47 of *Proc. Sympos. Pure Math.*, pages 181–210. Amer. Math. Soc., Providence, RI, 1987. With an appendix by S. P. Norton.

[McM92] Curt McMullen. Riemann surfaces and the geometrization of 3-manifolds. *Bull. Amer. Math. Soc. (N.S.)*, 27(2):207–216, 1992.

[MM07] Jan Manschot and Gregory Moore. A modern Farey tail. arXiv:0712.0573v1, December 2007.

[MSS09] Alexander Maloney, Wei Song, and Andrew Strominger. Chiral gravity, log gravity and extremal CFT. arXiv:0903.4573v1, March 2009.

[MW07] A. Maloney and E. Witten. Quantum gravity partition functions in three dimensions. arXiv:0712.0155v1, December 2007.

[Nie73] Douglas Niebur. A class of nonanalytic automorphic functions. *Nagoya Math. J.*, 52:133–145, 1973.

[Nie74] Douglas Niebur. Construction of automorphic forms and integrals. *Trans. Amer. Math. Soc.*, 191:373–385, 1974.

[Nor84] S. P. Norton. More on moonshine. In *Computational group theory (Durham, 1982)*, pages 185–193. Academic Press, London, 1984.

[Pri00] Wladimir de Azevedo Pribitkin. A generalization of the Goldfeld-Sarnak estimate on Selberg's Kloosterman zeta-function. *Forum Math.*, 12(4):449–459, 2000.

[Rad38] Hans Rademacher. The Fourier Coefficients of the Modular Invariant $\mathrm{J}(\tau)$. *Amer. J. Math.*, 60(2):501–512, 1938.

[Rad39] Hans Rademacher. The Fourier Series and the Functional Equation of the Absolute Modular Invariant $\mathrm{J}(\tau)$. *Amer. J. Math.*, 61(1):237–248, 1939.

[Sel65] Atle Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.

[Sul81] Dennis Sullivan. On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions. In *Riemann surfaces and related topics: Proceedings of the*

1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), volume 97 of *Ann. of Math. Stud.*, pages 465–496. Princeton Univ. Press, Princeton, N.J., 1981.

[Tho79a] J. G. Thompson. Finite groups and modular functions. *Bull. London Math. Soc.*, 11(3):347–351, 1979.

[Tho79b] J. G. Thompson. Some numerology between the Fischer-Griess Monster and the elliptic modular function. *Bull. London Math. Soc.*, 11(3):352–353, 1979.

[Tui95] Michael P. Tuite. On the relationship between Monstrous Moonshine and the uniqueness of the Moonshine module. *Comm. Math. Phys.*, 166(3):495–532, 1995.

[Tui08] Michael Tuite. Monstrous and Generalized Moonshine and permutation orbifolds. arXiv:0811.4525v1, November 2008.

[Wit87] Edward Witten. Physics and geometry. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, pages 267–303, Providence, RI, 1987. Amer. Math. Soc.

[Wit07] Edward Witten. Three-dimensional gravity revisited. arXiv:0706.3359v1, June 2007.

arXiv:1206.5143v1 [hep-th] 22 Jun 2012

AEI-2012-058

Mathieu Moonshine and Orbifold K3s[*]

Matthias R. Gaberdiel
Institut für Theoretische Physik
ETH Zurich, 8093 Zurich, Switzerland

and

Roberto Volpato
Max-Planck-Institut für Gravitationsphysik
Am Mühlenberg 1, 14476 Golm, Germany

Abstract: The current status of 'Mathieu Moonshine', the idea that the Mathieu group $\mathbb{M}_{24}$ organises the elliptic genus of K3, is reviewed. While there is a consistent decomposition of all Fourier coefficients of the elliptic genus in terms of Mathieu $\mathbb{M}_{24}$ representations, a conceptual understanding of this phenomenon in terms of K3 sigma-models is still missing. In particular, it follows from the recent classification of the automorphism groups of arbitrary K3 sigma-models that (i) there is no single K3 sigma-model that has $\mathbb{M}_{24}$ as an automorphism group; and (ii) there exist 'exceptional' K3 sigma-models whose automorphism group is not even a subgroup of $\mathbb{M}_{24}$. Here we show that all cyclic torus orbifolds are exceptional in this sense, and that almost all of the exceptional cases are realised as cyclic torus orbifolds. We also provide an explicit construction of a $\mathbb{Z}_5$ torus orbifold that realises one exceptional class of K3 sigma-models.

1 Introduction

In 2010, Eguchi, Ooguri and Tachikawa observed that the elliptic genus of K3 shows signs of an underlying Mathieu $\mathbb{M}_{24}$ group action [1]. In particular, they noted (see section 2 below for more details) that the Fourier coefficients of the elliptic genus can be written as sums of dimensions of irreducible $\mathbb{M}_{24}$ representations.[†] This intriguing observation is very reminiscent of the famous realisation of McKay and Thompson who noted that the Fourier expansion coefficients of the J-function can be written in terms of dimensions of representations of the Monster group [2, 3]. This led to a development that is now usually referred to as 'Monstrous Moonshine', see [4] for a nice review. One important upshot of that analysis was that the J-function can be thought of as the partition function of a self-dual conformal field theory, the 'Monster conformal field theory' [5, 6], whose automorphism group is precisely the Monster group. The existence of this conformal field theory explains many aspects of Monstrous Moonshine although not all — in particular, the genus zero property is rather mysterious from this point of view.

---

[*] Partially based on talk given by M.R.G. at the conference 'Conformal Field Theory, Automorphic Forms and Related Topics', Heidelberg, 19-23 September 2011.

[†] Actually, they did not just look at the Fourier coefficients themselves, but at the decomposition of the elliptic genus with respect to the elliptic genera of irreducible $\mathcal{N} = 4$ superconformal representations. They then noted that these expansion coefficients (and hence in particular the usual Fourier coefficients) are sums of dimensions of irreducible $\mathbb{M}_{24}$ representations.

In the Mathieu case, the situation is somewhat different compared to the early days of Monstrous Moonshine. It is by construction clear that the underlying conformal field theory *is* a K3 sigma-model (describing string propagation on the target space K3). However, this does not characterise the corresponding conformal field theory uniquely as there are many inequivalent such sigma-models — in fact, there is an 80-dimensional moduli space of such theories, all of which lead to the same elliptic genus. The natural analogue of the 'Monster conformal field theory' would therefore be a special K3 sigma-model whose automorphism group coincides with $\mathbb{M}_{24}$. Unfortunately, as we shall review here (see section 3), such a sigma-model does not exist: we have classified the automorphism groups of all K3 sigma-models, and none of them contains $\mathbb{M}_{24}$ [7]. In fact, not even all automorphism groups are contained in $\mathbb{M}_{24}$: the exceptional cases are the possibilities (ii), (iii) and (iv) of the theorem in section 3 (see [7]), as well as case (i) for nontrivial G'. Case (iii) was already shown in [7] to be realised by a specific Gepner model that is believed to be equivalent to a torus orbifold by $\mathbb{Z}_3$. Here we show that also cases (ii) and (iv) are realised by actual K3s — the argument in [7] for this relied on some assumption about the regularity of K3 sigma-models — and in both cases the relevant K3s are again torus orbifolds. More specifically, case (ii) is realised by an asymmetric $\mathbb{Z}_5$ orbifold of $\mathbb{T}^4$ (see section 5),[‡] while for case (iii) the relevant orbifold is by $\mathbb{Z}_3$ (see section 6).

Cyclic torus orbifolds are rather special K3s since they always possess a quantum symmetry whose orbifold leads back to $\mathbb{T}^4$. Using this property of cyclic torus orbifolds, we show (see section 4) that the group of automorphisms of K3s that are cyclic torus orbifolds is always exceptional; in particular, the quantum symmetry itself is never an element of $\mathbb{M}_{24}$. Although some 'exceptional' automorphism groups (contained in case (i) of the classification theorem) can also arise in K3 models that are not cyclic torus orbifolds, our observation may go a certain way towards explaining why only $\mathbb{M}_{24}$ seems to appear in the elliptic genus of K3.

We should mention that Mathieu Moonshine can also be formulated in terms of a mock modular form that can be naturally associated to the elliptic genus of K3 [1, 9, 10, 11]; this point of view has recently led to an interesting class of generalisations [12]. There are also indications that, just as for Monstrous Moonshine, Mathieu Moonshine can possibly be understood in terms of an underlying Borcherds-Kac-Moody algebra [13, 14, 15, 16].

2 Mathieu Moonshine

Let us first review the basic idea of 'Mathieu Moonshine'. We consider a conformal field theory sigma-model with target space K3. This theory has $\mathcal{N} = (4, 4)$ superconformal symmetry on the world-sheet. As a consequence, the space of states can be decomposed into representations of the $\mathcal{N} = 4$ superconformal algebra, both for the left- and the right-movers. (The left- and right-moving actions commute, and thus we can find a simultaneous decomposition.) The full space of states takes then the form

---

[‡] Since the orbifold action is asymmetric, this evades various no-go-theorems (see e.g. [8]) that state that the possible orbifold groups are either $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_4$, or $\mathbb{Z}_6$.

$$\mathcal{H} = \bigoplus_{i,j} N_{ij}\, \mathcal{H}_i \otimes \bar{\mathcal{H}}_j \ , \tag{2.1}$$

where i and j label the different $\mathcal{N} = 4$ superconformal representations, and $N_{ij} \in \mathbb{N}_0$ denote the multiplicities with which these representations appear. The $\mathcal{N} = 4$ algebra contains, apart from the Virasoro algebra L_n at $c = 6$, four supercharge generators, as well as an affine $\hat{\mathfrak{su}}(2)_1$ subalgebra at level one; we denote the Cartan generator of the zero mode subalgebra $\mathfrak{su}(2)$ by J_0.

The full partition function of the conformal field theory is quite complicated, and is only explicitly known at special points in the moduli space. However, there exists some sort of partial index that is much better behaved. This is the so-called *elliptic genus* that is defined by

$$\phi_{\rm K3}(\tau, z) = \mathrm{Tr}_{\rm RR}\Big(q^{L_0-\frac{c}{24}}\, y^{J_0}\, (-1)^F\, \bar{q}^{\bar{L}_0-\frac{\bar{c}}{24}}\, (-1)^{\bar{F}}\Big) \equiv \phi_{0,1}(\tau, z) \ . \tag{2.2}$$

Here the trace is only taken over the Ramond-Ramond part of the spectrum (2.1), and the right-moving $\mathcal{N} = 4$ modes are denoted by a bar. Furthermore, $q = \exp(2\pi i\tau)$ and $y = \exp(2\pi i z)$, F and $\bar{F}$ are the left- and right-moving fermion number operators, and the two central charges equal $c = \bar{c} = 6$. Note that the elliptic genus does not actually depend on $\bar{\tau}$, although $\bar{q} = \exp(-2\pi i\bar{\tau})$ does; the reason for this is that, with respect to the right-moving algebra, the elliptic genus is like a Witten index, and only the right-moving ground states contribute. To see this one notices that states that are not annihilated by a supercharge zero mode appear always as a boson-fermion pair; the contribution of such a pair to the elliptic genus however vanishes because the two states contribute with the opposite sign (as a consequence of the $(-1)^{\bar{F}}$ factor). Thus only the right-moving ground states, i.e. the states that are annihilated by all right-moving supercharge zero modes, contribute to the elliptic genus, and the commutation relations of the $\mathcal{N} = 4$ algebra then imply that they satisfy $(\bar{L}_0 - \frac{\bar{c}}{24})\phi_{\rm ground} = 0$; thus it follows that the elliptic genus is independent of $\bar{\tau}$. Note that this argument does not apply to the left-moving contributions because of the y^{J_0} factor. (The supercharges are 'charged' with respect to the J_0 Cartan generator, and hence the two terms of a boson-fermion pair come with different powers of y. However, if we also set $y = 1$, the elliptic genus does indeed become a constant, independent of τ and $\bar{\tau}$.)

It follows from general string considerations that the elliptic genus defines a *weak Jacobi form of weight zero and index one* [17]. Recall that a weak Jacobi form of weight w and index m is a function [18]

$$\phi_{w,m} : \mathbb{H}_+ \times \mathbb{C} \to \mathbb{C} \ , \qquad (\tau, z) \mapsto \phi_{w,m}(\tau, z) \tag{2.3}$$

that satisfies

$$\phi_{w,m}\Big(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d}\Big) = (c\tau + d)^w\, e^{2\pi i m \frac{cz^2}{c\tau+d}}\, \phi_{w,m}(\tau, z) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \ , \tag{2.4}$$

$$\phi(\tau, z + \ell\tau + \ell') = e^{-2\pi i m(\ell^2\tau + 2\ell z)}\phi(\tau, z) \qquad\qquad \ell, \ell' \in \mathbb{Z} \ , \tag{2.5}$$

and has a Fourier expansion

$$\phi(\tau,z) = \sum_{n\geq 0,\, \ell\in\mathbb{Z}} c(n,\ell) q^n y^\ell \tag{2.6}$$

with $c(n,\ell) = (-1)^w c(n,-\ell)$. Weak Jacobi forms have been classified, and there is only one weak Jacobi form with $w=0$ and $m=1$. Up to normalisation $\phi_{\rm K3}$ must therefore agree with this unique weak Jacobi form $\phi_{0,1}$, which can explicitly be written in terms of Jacobi theta functions as

$$\phi_{0,1}(\tau,z) = 8 \sum_{i=2,3,4} \frac{\vartheta_i(\tau,z)^2}{\vartheta_i(\tau,0)^2} \ . \tag{2.7}$$

Note that the Fourier coefficients of $\phi_{\rm K3}$ are integers; as a consequence they cannot change continuously as one moves around in the moduli space of K3 sigma-models, and thus $\phi_{\rm K3}$ must be actually independent of the specific K3 sigma-model that is being considered, i.e. independent of the point in the moduli space. Here we have used that the moduli space is connected. More concretely, it can be described as the double quotient

$$\mathcal{M}_{\rm K3} = {\rm O}(\Gamma^{4,20}) \setminus {\rm O}(4,20) \ / \ {\rm O}(4)\times {\rm O}(20) \ . \tag{2.8}$$

We can think of the Grassmannian on the right

$${\rm O}(4,20) \ / \ {\rm O}(4)\times {\rm O}(20) \tag{2.9}$$

as describing the choice of a positive-definite 4-dimensional subspace $\Pi \subset \mathbb{R}^{4,20}$, while the group on the left, ${\rm O}(\Gamma^{4,20})$, leads to discrete identifications among them. Here ${\rm O}(\Gamma^{4,20})$ is the group of isometries of a given fixed unimodular lattice $\Gamma^{4,20} \subset \mathbb{R}^{4,20}$. (In physics terms, the lattice $\Gamma^{4,20}$ can be thought of as the D-brane charge lattice of string theory on K3.)

Let us denote by $\mathcal{H}^{(0)} \subset \mathcal{H}_{\rm RR}$ the subspace of (2.1) that consists of those RR states for which the right-moving states are ground states. (Thus $\mathcal{H}^{(0)}$ consists of the states that contribute to the elliptic genus.) $\mathcal{H}^{(0)}$ carries an action of the left-moving $\mathcal{N}=4$ superconformal algebra, and at any point in moduli space, its decomposition is of the form

$$\mathcal{H}^{(0)} = 20\cdot \mathcal{H}_{h=\frac{1}{4},j=0} \ \oplus \ 2\cdot \mathcal{H}_{h=\frac{1}{4},j=\frac{1}{2}} \ \oplus \ \bigoplus_{n=1}^{\infty} D_n\, \mathcal{H}_{h=\frac{1}{4}+n,j=\frac{1}{2}} \ , \tag{2.10}$$

where $\mathcal{H}_{h,j}$ denotes the irreducible $\mathcal{N}=4$ representation whose Virasoro primary states have conformal dimension h and transform in the spin j representation of $\mathfrak{su}(2)$. The multiplicities D_n are *not* constant over the moduli space, but the above argument shows that

$$A_n = {\rm Tr}_{D_n}(-1)^{\bar F} \tag{2.11}$$

are (where D_n is now understood not just as a multiplicity, but as a representation of the right-moving $(-1)^{\bar F}$ operator that determines the sign with which these states contribute to the elliptic genus). In this language, the elliptic genus then takes the form

$$\phi_{\rm K3}(\tau,z) = 20\cdot \chi_{h=\frac{1}{4},j=0}(\tau,z) - 2\cdot \chi_{h=\frac{1}{4},j=\frac{1}{2}}(\tau,z) + \sum_{n=1}^{\infty} A_n \cdot \chi_{h=\frac{1}{4}+n,j=\frac{1}{2}}(\tau,z) \ , \tag{2.12}$$

where $\chi_{h,j}(\tau,z)$ is the 'elliptic' genus of the corresponding $\mathcal{N}=4$ representation,

$$\chi_{h,j}(\tau,z) = \mathrm{Tr}_{\mathcal{H}_{h,j}}\Big(q^{L_0-\frac{c}{24}}\, y^{J_0}(-1)^F\Big) \ , \tag{2.13}$$

and we have used that $(-1)^{\bar{F}}$ takes the eigenvalues $+1$ and -1 on the 20- and 2-dimensional multiplicity spaces of the first two terms in (2.10), respectively.

The key observation of Eguchi, Ooguri & Tachikawa (EOT) [1] was that the A_n are sums of dimensions of $\mathbb{M}_{24}$ representation, in striking analogy to the original Monstrous Moonshine conjecture of [3]; the first few terms are

$$\begin{aligned} A_1 &= 90 = \mathbf{45} + \overline{\mathbf{45}} && (2.14)\\ A_2 &= 462 = \mathbf{231} + \overline{\mathbf{231}} && (2.15)\\ A_3 &= 1540 = \mathbf{770} + \overline{\mathbf{770}} \ , && (2.16) \end{aligned}$$

where $\mathbf{N}$ denotes a representation of $\mathbb{M}_{24}$ of dimension N. Actually, they guessed correctly the first six coefficients; from A_7 onwards the guesses become much more ambiguous (since the dimensions of the $\mathbb{M}_{24}$ representations are not that large) and they actually misidentified the seventh coefficient in their original analysis. (We will come back to the question of why and how one can be certain about the 'correct' decomposition shortly, see section 2.2.) The alert reader will also notice that the first two coefficients in (2.10), namely 20 and -2, are not directly $\mathbb{M}_{24}$ representations; the correct prescription is to introduce virtual representations and to write

$$20 = \mathbf{23} - 3\cdot\mathbf{1} \ , \qquad -2 = -2\cdot\mathbf{1} \ . \tag{2.17}$$

Recall that $\mathbb{M}_{24}$ is a sporadic finite simple group of order

$$|\mathbb{M}_{24}| = 2^{10}\cdot 3^3\cdot 5\cdot 7\cdot 11\cdot 23 = 244\,823\,040 \ . \tag{2.18}$$

It has 26 conjugacy classes (which are denoted by 1A, 2A, 3A, …, 23A, 23B, where the number refers to the order of the corresponding group element) — see eqs. (2.19) and (2.20) below for the full list — and therefore also 26 irreducible representations whose dimensions range from $N=1$ to $N=10395$. The Mathieu group $\mathbb{M}_{24}$ can be defined as the subgroup of the permutation group S_{24} that leaves the extended Golay code invariant; equivalently, it is the quotient of the automorphism group of the $\mathfrak{su}(2)^{24}$ Niemeier lattice, divided by the Weyl group. Thought of as a subgroup of $\mathbb{M}_{24}\subset S_{24}$, it contains the subgroup $\mathbb{M}_{23}$ that is characterised by the condition that it leaves a given (fixed) element of $\{1,\ldots,24\}$ invariant.

2.1 Classical symmetries

The appearance of a Mathieu group in the elliptic genus of K3 is not totally surprising in view of the Mukai theorem [19, 20]. It states that any finite group of symplectic automorphisms of a K3 surface can be embedded into the Mathieu group $\mathbb{M}_{23}$. The symplectic automorphisms of a K3 surface define symmetries that act on the multiplicity spaces of the $\mathcal{N}=4$ representations, and therefore explain part of the above findings. However, it is also clear from Mukai's argument that they do not even account for the full $\mathbb{M}_{23}$ group. Indeed, every symplectomorphism of K3 has at least five orbits on the set $\{1,\ldots,24\}$, and thus not all elements of $\mathbb{M}_{23}$ can be realised as a symplectomorphism. More specifically, of the 26 conjugacy classes of $\mathbb{M}_{24}$, 16 have a representative in $\mathbb{M}_{23}$, namely

$$\text{repr. in } \mathbb{M}_{23}: \quad \begin{array}{ll} \text{1A, 2A, 3A, 4B, 5A, 6A, 7A, 7B, 8A} & \text{(geometric)} \\ \text{11A, 14A, 14B, 15A, 15B, 23A, 23B} & \text{(non-geometric)}\ , \end{array} \tag{2.19}$$

where 'geometric' means that a representative can be (and in fact is) realised by a geometric symplectomorphism (i.e. that the representative has at least five orbits when acting on the set $\{1,\ldots,24\}$), while 'non-geometric' means that this is not the case. The remaining conjugacy classes do *not* have a representative in $\mathbb{M}_{23}$, and are therefore not accounted for geometrically via the Mukai theorem

$$\text{no repr. in } \mathbb{M}_{23}: \quad \text{2B, 3B, 4A, 4C, 6B, 10A, 12A, 12B, 21A, 21B}\ . \tag{2.20}$$

The classical symmetries can therefore only explain the symmetries in the first line of (2.19). Thus an additional argument is needed in order to understand the origin of the other symmetries; we shall come back to this in section 3.

2.2 Evidence for Moonshine

As was already alluded to above, in order to determine the 'correct' decomposition of the A_n multiplicity spaces in terms of $\mathbb{M}_{24}$ representations, we need to study more than just the usual elliptic genus. By analogy with Monstrous Moonshine, the natural objects to consider are the analogues of the McKay Thompson series [21]. These are obtained from the elliptic genus upon replacing

$$A_n = \dim R_n \ \rightarrow\ \mathrm{Tr}_{R_n}(g)\ , \tag{2.21}$$

where $g \in \mathbb{M}_{24}$, and R_n is the $\mathbb{M}_{24}$ representation whose dimension equals the coefficient A_n; the resulting functions are then (compare (2.12))

$$\phi_g(\tau,z) = \mathrm{Tr}_{\mathbf{23}-3\cdot\mathbf{1}}(g)\,\chi_{h=\frac{1}{4},j=0}(\tau,z) - 2\,\mathrm{Tr}_{\mathbf{1}}(g)\,\chi_{h=\frac{1}{4},j=\frac{1}{2}}(\tau,z) + \sum_{n=1}^{\infty} \mathrm{Tr}_{R_n}(g)\,\chi_{h=\frac{1}{4}+n,j=\frac{1}{2}}(\tau,z)\ . \tag{2.22}$$

The motivation for this definition comes from the observation that if the underlying vector space $\mathcal{H}^{(0)}$, see eq. (2.10), of states contributing to the elliptic genus were to carry an action of $\mathbb{M}_{24}$, $\phi_g(\tau,z)$ would equal the 'twining elliptic genus', i.e. the elliptic genus twined by the action of g

$$\phi_g(\tau,z) = \mathrm{Tr}_{\mathcal{H}^{(0)}}\Big(g\, q^{L_0-\frac{c}{24}}\, y^{J_0}\, (-1)^F\, \bar{q}^{\bar{L}_0-\frac{\bar{c}}{24}}\, (-1)^{\bar{F}}\Big)\ . \tag{2.23}$$

Obviously, a priori, it is not clear what the relevant R_n in (2.21) are. However, we have some partial information about them:

(i) For any explicit realisation of a symmetry of a K3 sigma-model, we can calculate (2.23) directly. (In particular, for some symmetries, the relevant twining genera had already been calculated in [22].)

(ii) The observation of EOT determines the first six coefficients explicitly.

(iii) The twining genera must have special modular properties.

Let us elaborate on (iii). Assuming that the functions $\phi_g(\tau, z)$ have indeed an interpretation as in (2.23), they correspond in the usual orbifold notation of string theory to the contribution

$$\phi_g(\tau, z) \;\longleftrightarrow\; e\ \underset{g}{\boxed{}} \tag{2.24}$$

where e is the identity element of the group. Under a modular transformation it is believed that these twining and twisted genera transform (up to a possible phase) as

$$h\ \underset{g}{\boxed{}} \quad \xrightarrow{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \quad h^d g^c\ \underset{g^a h^b}{\boxed{}} \tag{2.25}$$

The twining genera (2.24) are therefore invariant (possibly up to a phase) under the modular transformations with

$$\gcd(a, o(g)) = 1 \qquad \text{and} \qquad c = 0 \mod o(g), \tag{2.26}$$

where $o(g)$ is the order of the group element g and we used that for $\gcd(a, o(g)) = 1$, the group element g^a is in the same conjugacy class as g or g^{-1}. (Because of reality, the twining genus of g and g^{-1} should be the same.) Since $ad - bc = 1$, the second condition implies the first, and we thus conclude that $\phi_g(\tau, z)$ should be (up to a possible multiplier system) a weak Jacobi form of weight zero and index one under the subgroup of $\mathrm{SL}(2, \mathbb{Z})$

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \ : \ c = 0 \bmod (N) \right\} , \tag{2.27}$$

where $N = o(g)$. This is a relatively strong condition, and knowing the first few terms (for a fixed multiplier system) determines the function uniquely. In order to use this constraint, however, it is important to know the multiplier system. An ansatz (that seems to work, see below) was made in [23]

$$\phi_g\Big(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d}\Big) = e^{\frac{2\pi i c d}{N h}}\, e^{\frac{2\pi i\, c z^2}{c\tau + d}}\, \phi_g(\tau, z) \ , \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) \ , \tag{2.28}$$

where N is again the order of g and $h | \gcd(N, 12)$. The multiplier system is trivial ($h = 1$) if and only if g contains a representative in $\mathbb{M}_{23} \subset \mathbb{M}_{24}$. For the other conjugacy classes, the values are tabulated in table 1. It was noted in [10] that h equals the length of the shortest cycle (when interpreted as a permutation in S_{24}, see table 1 of [10]).

Using this ansatz, explicit expressions for all twining genera were determined in [23]; independently, the same twining genera were also found (using guesses based on the cycle shapes of the corresponding S_{24} representations) in [24]. (Earlier partial results had been obtained in [9] and [25].)

Class	2B	3B	4A	4C	6B	10A	12A	12B	21AB
h	2	3	2	4	6	2	2	12	3

Table 1: Value of h for the conjugacy classes in (2.28).

These explicit expressions for the twining genera then allow for a very non-trivial check of the EOT proposal. As is clear from their definition in (2.22), they determine the coefficients

$$\mathrm{Tr}_{R_n}(g) \qquad \text{for all } g \in \mathbb{M}_{24} \text{ and all } n \geq 1. \tag{2.29}$$

This information is therefore sufficient to *determine* the representations R_n, i.e. to calculate their decomposition into irreducible $\mathbb{M}_{24}$ representations, for all n. We have worked out the decompositions explicitly for the first 500 coefficients, and we have found that each R_n can be written as a direct sum of $\mathbb{M}_{24}$ representations with non-negative integer multiplicities [23]. (Subsequently [24] tested this property for the first 600 coefficients, and apparently Tachikawa has also checked it for the first 1000 coefficients.) Terry Gannon has informed us that this information is sufficient to prove that the same will then happen for all n [26]. In some sense this then proves the EOT conjecture.

3 Symmetries of K3 models

While the above considerations establish in some sense the EOT conjecture, they do not offer any insight into why the elliptic genus of K3 should exhibit an $\mathbb{M}_{24}$ symmetry. This is somewhat similar to the original situation in Monstrous Moonshine, after Conway and Norton had found the various Hauptmodules by somewhat similar techniques. Obviously, in the case of Monstrous Moonshine, many of these observations were subsequently explained by the construction of the Monster CFT (that possesses the Monster group as its automorphism group) [5, 6]. So we should similarly ask for a microscopic explanation of these findings.

In some sense it is clear what the analogue of the Monster CFT in the current context should be: we know that the function in question *is* the elliptic genus of K3. However, there is one problem with this. As we mentioned before, there is not just one K3 sigma-model, but rather a whole moduli space (see eq. (2.8)) of such CFTs. So the simplest explanation of the EOT observation would be if there is (at least) one special K3 sigma-model that has $\mathbb{M}_{24}$ as its automorphism group. Actually, the relevant symmetry group should commute with the action of the $\mathcal{N} = (4,4)$ superconformal symmetry (since it should act on the multiplicity spaces in $\mathcal{H}^{(0)}$, see eq. (2.10)). Furthermore, since the two $\mathcal{N} = 4$ representations with $h = \frac{1}{4}$ and $j = \frac{1}{2}$ are singlets — recall that the coefficient -2 transforms as $-2 = -2 \cdot \mathbf{1}$, see (2.17) — the automorphism must act trivially on the 4 RR ground states that transform in the $(\mathbf{2},\mathbf{2})$ representation of the $\mathfrak{su}(2)_L \times \mathfrak{su}(2)_R$ subalgebra of $\mathcal{N} = (4,4)$. Note that these four states generate the simultaneous half-unit spectral flows in the left- and the right-moving sector; the requirement that the symmetry leaves them invariant therefore means that spacetime supersymmetry is preserved.

Recall from (2.8) that the different K3 sigma-models are parametrised by the choice of a positive-definite 4-dimensional subspace $\Pi \subset \mathbb{R}^{4,20}$, modulo some discrete identifications. Let us denote by G_Π the group of symmetries of the sigma-model described by Π that commute with the action of $\mathcal{N} = (4,4)$ and preserve the RR ground states in the $(\mathbf{2},\mathbf{2})$ (see above). It was argued in [7] that G_Π is precisely the subgroup of $\mathrm{O}(\Gamma^{4,20})$ consisting of those elements that leave Π pointwised fixed. The possible symmetry groups G_Π can then be classified following essentially the paradigm of the Mukai-Kondo argument for the symplectomorphisms of K3 surfaces [19, 20]. The outcome of the analysis can be summarised by the following theorem [7]:

Theorem: *Let G be the group of symmetries of a non-linear sigma-model on $K3$ preserving the $\mathcal{N} = (4,4)$ superconformal algebra as well as the spectral flow operators. One of the following possibilities holds:*

(i) $G = G'.G''$, *where G' is a subgroup of $\mathbb{Z}_2^{11}$, and G'' is a subgroup of $\mathbb{M}_{24}$ with at least four orbits when acting as a permutation on* $\{1,\ldots,24\}$

(ii) $G = 5^{1+2} : \mathbb{Z}_4$

(iii) $G = \mathbb{Z}_3^4 : A_6$

(iv) $G = 3^{1+4} : \mathbb{Z}_2.G''$, *where G'' is either trivial, $\mathbb{Z}_2$ or $\mathbb{Z}_2^2$.*

Here $G = N.Q$ means that N is a normal subgroup of G, and $G/N \cong Q$; when G is the semidirect product of N and Q, we denote it by $N : Q$. Furthermore, p^{1+2n} is an extra-special group of order p^{1+2n}, which is an extension of $\mathbb{Z}_p^{2n}$ by a central element of order p.

We will give a sketch of the proof below (see section 3.1), but for the moment let us comment on the implications of this result. First of all, our initial expectation from above is not realised: none of these groups $G \equiv G_\Pi$ contains $\mathbb{M}_{24}$. In particular, the twining genera for the conjugacy classes 12B, 21A, 21B, 23A, 23B of $\mathbb{M}_{24}$ cannot be realised by any symmetry of a K3 sigma-model. Thus we cannot give a direct explanation of the EOT observation along these lines.

Given that the elliptic genus is constant over the moduli space, one may then hope that we can explain the origin of $\mathbb{M}_{24}$ by 'combining' symmetries from different points in the moduli space. As we have mentioned before, this is also similar to what happens for the geometric symplectomorphisms of K3: it follows from the Mukai theorem that the Mathieu group $\mathbb{M}_{23}$ is the smallest group that contains all symplectomorphisms, but there is no K3 surface where all of $\mathbb{M}_{23}$ is realised, and indeed, certain generators of $\mathbb{M}_{23}$ can never be symmetries, see (2.19). However, also this explanation of the EOT observation is somewhat problematic: as is clear from the above theorem, not all symmetry groups of K3 sigma-models are in fact subgroups of $\mathbb{M}_{24}$. In particular, none of the cases (ii), (iii) and (iv) (as well as case (i) with G' non-trivial) have this property, as can be easily seen by comparing the prime factor decompositions of their orders to (2.18). The smallest group that contains all groups of the theorem is the Conway group Co_1, but as far as we are aware, there is no evidence of any 'Conway Moonshine' in the elliptic genus of K3.

One might speculate that, generically, the group G must be a subgroup of $\mathbb{M}_{24}$, and that the models whose symmetry group is not contained in $\mathbb{M}_{24}$ are, in some sense, special or 'exceptional' points in the moduli space. In order to make this idea precise, it is useful to analyse the exceptional models in detail. In [7], some examples have been provided of case (i) with non-trivial G' (a torus orbifold $\mathbb{T}^4/\mathbb{Z}_2$ or the Gepner model 2^4, believed to be equivalent to a $\mathbb{T}^4/\mathbb{Z}_4$ orbifold), and of case (iii) (the Gepner model 1^6, which is believed to be equivalent to a $\mathbb{T}^4/\mathbb{Z}_3$ orbifold, see also [27]). For the cases (ii) and (iv), only an existence proof was given. In section 5, we will improve the situation by constructing in detail an example of case (ii), realised as an asymmetric $\mathbb{Z}_5$-orbifold of a torus $\mathbb{T}^4$. Furthermore, in section 6 we will briefly discuss the $\mathbb{Z}_3$-orbifold of a torus and the explicit realisation of its symmetry group, corresponding to cases (ii) and (iv) for any G''.

Notice that all the examples of exceptional models known so far are provided by torus orbifolds. In fact, we will show below (see section 4) that all cyclic torus orbifolds have exceptional symmetry groups. Conversely, we will prove that the cases (ii)–(iv) of the theorem are always realised by (cyclic) torus orbifolds. On the other hand, as we shall also explain, some of the exceptional models in case (i) are not cyclic torus orbifolds.

3.1 Sketch of the proof of the Theorem

In this subsection, we will describe the main steps in the proof of the above theorem; the details can be found in [7].

It was argued in [7] that the supersymmetry preserving automorphisms of the non-linear sigma-model characterised by Π generate the group $G \equiv G_\Pi$ that consists of those elements of $O(\Gamma^{4,20})$ that leave Π pointwise fixed. Let us denote by L^G the sublattice of G-invariant vectors of $L \equiv \Gamma^{4,20}$, and define L_G to be its orthogonal complement that carries a genuine action of G. Since $L^G \otimes \mathbb{R}$ contains the subspace Π, it follows that L^G has signature $(4, d)$ for some $d \geq 0$, so that L_G is a negative-definite lattice of rank $20 - d$. In [7], it is proved that, for any consistent model, L_G can be embedded (up to a change of sign in its quadratic form) into the Leech lattice Λ, the unique 24-dimensional positive-definite even unimodular lattice containing no vectors of squared norm 2. Furthermore, the action of G on L_G can be extended to an action on the whole of Λ, such that the sublattice $\Lambda^G \subset \Lambda$ of vectors fixed by G is the orthogonal complement of L_G in Λ. This construction implies that G must be a subgroup of $\mathrm{Co}_0 \equiv \mathrm{Aut}(\Lambda)$ that fixes a sublattice Λ^G of rank $4 + d$. Conversely, it can be shown that any such subgroup of $\mathrm{Aut}(\Lambda)$ is the symmetry group of some K3 sigma-model.

This leaves us with characterising the possible subgroups of the finite group $\mathrm{Co}_0 \equiv \mathrm{Aut}(\Lambda)$ that stabilise a suitable sublattice; problems of this kind have been studied in the mathematical literature before. In particular, the stabilisers of sublattices of rank at least 4 are, generically, the subgroups of $\mathbb{Z}_2^{11} : \mathbb{M}_{24}$ described in case (i) of the theorem above. The three cases (ii), (iii), (iv) arise when the invariant sublattice Λ^G is contained in some $\mathcal{S}$-lattice $S \subset \Lambda$. An $\mathcal{S}$-lattice S is a primitive sublattice of Λ such that each vector of S is congruent modulo $2S$ to a vector of norm 0, 4 or 6. Up to isomorphisms, there are only three kind of $\mathcal{S}$-lattices of rank at least four; their properties are summarised in the following table:

Name	type	rk S	Stab(S)	Aut(S)
$(A_2 \oplus A_2)'(3)$	$2^9 3^6$	4	$\mathbb{Z}_3^4 : A_6$	$\mathbb{Z}_2 \times (S_3 \times S_3).\mathbb{Z}_2$
$A_4^*(5)$	$2^5 3^{10}$	4	$5^{1+2} : \mathbb{Z}_4$	$\mathbb{Z}_2 \times S_5$
$E_6^*(3)$	$2^{27} 3^{36}$	6	$3^{1+4} : \mathbb{Z}_2$	$\mathbb{Z}_2 \times W(E_6)$.

Here, S is characterised by the type $2^p 3^q$, which indicates that S contains p pairs of opposite vectors of norm 4 (type 2) and q pairs of opposite vectors of norm 6 (type 3).

The group $\mathrm{Stab}(S)$ is the pointwise stabiliser of S in Co_0 and $\mathrm{Aut}(S)$ is the quotient of the setwise stabiliser of S modulo its pointwise stabiliser $\mathrm{Stab}(S)$. The group $\mathrm{Aut}(S)$ always contains a central $\mathbb{Z}_2$ subgroup, generated by the transformation that inverts the sign of all vectors of the Leech lattice. The lattice of type $2^{27}3^{36}$ is isomorphic to the weight lattice (the dual of the root lattice) of E_6 with quadratic form rescaled by 3 (i.e. the roots in $E_6^*(3)$ have squared norm 6), and $\mathrm{Aut}(S)/\mathbb{Z}_2$ is isomorphic to the Weyl group $W(E_6)$ of E_6. Similarly, the lattice of type $2^5 3^{10}$ is the weight lattice of A_4 rescaled by 5, and $\mathrm{Aut}(S)/\mathbb{Z}_2$ is isomorphic to the Weyl group $W(A_4) \cong S_5$ of A_4. Finally, the type $2^9 3^6$ is the 3-rescaling of a lattice $(A_2 \oplus A_2)'$ obtained by adjoining to the root lattice $A_2 \oplus A_2$ an element $(e_1^*, e_2^*) \in A_2^* \oplus A_2^*$, with e_1^* and e_2^* of norm $2/3$. The latter $\mathcal{S}$-lattice can also be described as the sublattice of vectors of $E_6^*(3)$ that are orthogonal to an $A_2(3)$ sublattice of $E_6^*(3)$. The group $\mathrm{Aut}(S)/\mathbb{Z}_2$ is the product $(S_3 \times S_3).\mathbb{Z}_2$ of the Weyl groups $W(A_2) = S_3$, and the $\mathbb{Z}_2$ symmetry that exchanges the two A_2 and maps e_1^* to e_2^*.

The cases (ii)–(iv) of the above theorem correspond to Λ^G being isomorphic to $A_4^*(5)$ (case ii), to $(A_2 \oplus A_2)'(3)$ (case iii) or to a sublattice of $E_6^*(3)$ different from $(A_2 \oplus A_2)'(3)$ (case iv). In the cases (ii) and (iii), G is isomorphic to $\mathrm{Stab}(S)$. In case (iv), $\mathrm{Stab}(S)$ is, generically, a normal subgroup of G, and $G'' \cong G/\mathrm{Stab}(S)$ is a subgroup of $\mathrm{Aut}(S) \cong \mathbb{Z}_2 \times W(E_6)$ that fixes a sublattice $\Lambda^G \subseteq E_6^*(3)$, with $\Lambda^G \neq (A_2 \oplus A_2)'(3)$, of rank at least 4. The only non-trivial subgroups of $\mathbb{Z}_2 \times W(E_6)$ with these properties are $G'' = \mathbb{Z}_2$, which corresponds to Λ^G being orthogonal to a single vector of norm 6 in $E_6^*(3)$ (a rescaled root), and $G'' = \mathbb{Z}_2^2$, which corresponds to Λ^G being orthogonal to two orthogonal vectors of norm 6.[§] If Λ^G is orthogonal to two vectors $v_1, v_2 \in E_6^*(3)$ of norm 6, with $v_1 \cdot v_2 = -3$, then $\Lambda^G \cong (A_2 \oplus A_2)'(3)$ and case (iii) applies.

4 Symmetry groups of torus orbifolds

In this section we will prove that all K3 sigma-models that are realised as (possibly left-right asymmetric) orbifolds of $\mathbb{T}^4$ by a cyclic group have an 'exceptional' group of symmetries, i.e. their symmetries are not a subgroup of $\mathbb{M}_{24}$. Furthermore, these torus orbifolds account for most of the exceptional models (in particular, for all models in the cases (ii)–(iv) of the theorem). On the other hand, as we shall also explain, there are exceptional models in case (i) that are not cyclic torus orbifolds.

Our reasoning is somewhat reminiscent of the construction of [28, 29] in the context of Monstrous Moonshine. Any $\mathbb{Z}_n$-orbifold of a conformal field theory has an automorphism g of order n, called the *quantum symmetry*, which acts trivially on the untwisted sector and by multiplication by the phase $\exp(\frac{2\pi i k}{n})$ on the k-th twisted sector. Furthermore, the orbifold of the orbifold theory by the group generated by the quantum symmetry g, is equivalent to the original conformal field theory [30]. This observation is the key for characterising K3 models that can be realised as torus orbifolds:

A K3 model $\mathcal{C}$ is a $\mathbb{Z}_n$-orbifold of a torus model if and only if it has a symmetry g of order n such that $\mathcal{C}/\langle g\rangle$ is a consistent orbifold equivalent to a torus model.

[§] The possibility $G'' = \mathbb{Z}_4$ that has been considered in [7] has to be excluded, since there are no elements of order 4 in $W(E_6)$ that preserve a four-dimensional sublattice of $E_6^*(3)$.

In order to see this, suppose that $\mathcal{C}_{K3}$ is a K3 sigma-model that can be realised as a torus orbifold $\mathcal{C}_{K3} = \tilde{\mathcal{C}}_{\mathbb{T}^4}/\langle \tilde{g} \rangle$, where $\tilde{g}$ is a symmetry of order n of the torus model $\tilde{\mathcal{C}}_{\mathbb{T}^4}$. Then $\mathcal{C}_{K3}$ possesses a 'quantum symmetry' g of order n, such that the orbifold of $\mathcal{C}_{K3}$ by g describes again the original torus model, $\tilde{\mathcal{C}}_{\mathbb{T}^4} = \mathcal{C}_{K3}/\langle g \rangle$.

Conversely, suppose $\mathcal{C}_{K3}$ has a symmetry g of order n, such that the orbifold of $\mathcal{C}_{K3}$ by g is consistent, i.e. satisfies the level matching condition — this is the case if and only if the twining genus ϕ_g has a trivial multiplier system — and leads to a torus model $\mathcal{C}_{K3}/\langle g \rangle = \tilde{\mathcal{C}}_{\mathbb{T}^4}$. Then $\mathcal{C}_{K3}$ itself is a torus orbifold since we can take the orbifold of $\tilde{\mathcal{C}}_{\mathbb{T}^4}$ by the quantum symmetry associated to g, and this will, by construction, lead back to $\mathcal{C}_{K3}$.

Thus we conclude that $\mathcal{C} \equiv \mathcal{C}_{K3}$ can be realised as a torus orbifold if and only if $\mathcal{C}$ contains a symmetry g such that (i) ϕ_g has a trivial multiplier system; and (ii) the orbifold of $\mathcal{C}$ by g leads to a torus model $\tilde{\mathcal{C}}_{\mathbb{T}^4}$. It is believed that the orbifold of $\mathcal{C}$ by any $\mathcal{N} = (4,4)$-preserving symmetry group, if consistent, will describe a sigma-model with target space either a torus $\mathbb{T}^4$ or a K3 manifold. The two cases can be distinguished by calculating the elliptic genus; in particular, if the target space is a torus, the elliptic genus vanishes. Actually, since the space of weak Jacobi forms of weight zero and index one is 1-dimensional, this condition is equivalent to the requirement that the elliptic genus $\tilde{\phi}(\tau, z)$ of $\tilde{\mathcal{C}} = \mathcal{C}/\langle g \rangle$ vanishes at $z = 0$.

Next we recall that the elliptic genus of the orbifold by a group element g of order $n = o(g)$ is given by the usual orbifold formula

$$\tilde{\phi}(\tau, z) = \frac{1}{n} \sum_{i,j=1}^{n} \phi_{g^i, g^j}(\tau, z) \ , \tag{4.1}$$

where $\phi_{g^i,g^j}(\tau, z)$ is the twining genus for g^j in the g^i-twisted sector; this can be obtained by a modular transformation from the untwisted twining genus $\phi_{g^d}(\tau, z)$ with $d = \gcd(i, j, n)$. As we have explained above, it is enough to evaluate the elliptic genus for $z = 0$. Then

$$\phi_{g^d}(\tau, z = 0) = \mathrm{Tr}_{\mathbf{24}}(g^d) \ , \tag{4.2}$$

where $\mathrm{Tr}_{\mathbf{24}}(g^d)$ is the trace of g^d over the 24-dimensional space of RR ground states, and since (4.2) is constant (and hence modular invariant) we conclude that

$$\tilde{\phi}(\tau, 0) = \frac{1}{n} \sum_{i,j=1}^{n} \mathrm{Tr}_{\mathbf{24}}(g^{\gcd(i,j,n)}) \ . \tag{4.3}$$

According to the theorem in section 3, all symmetry groups of K3 sigma-models are subgroups of Co_0 and, in fact, $\mathrm{Tr}_{\mathbf{24}}(g^d)$ coincides with the trace of $g^d \in \mathrm{Co}_0$ in the standard 24-dimensional representation of Co_0. Thus, the elliptic genus of the orbifold model $\tilde{\mathcal{C}} = \mathcal{C}/\langle g \rangle$ only depends on the conjugacy class of g in Co_0. The group Co_0 contains 167 conjugacy classes, but only 42 of them contain symmetries that are realised by some K3 sigma-model, i.e. elements that fix at least a four-dimensional subspace in the standard 24-dimensional representation of Co_0. If $\mathrm{Tr}_{\mathbf{24}}(g) \neq 0$ (this happens for 31 of the above 42 conjugacy classes), the twining genus $\phi_g(\tau, z)$ has necessarily a trivial multiplier system, and the orbifold $\mathcal{C}/\langle g \rangle$ is consistent. These classes are listed in the following table, together with the dimension of the space that is fixed by g, the trace over the 24-dimensional representation, and the elliptic genus $\tilde{\phi}(\tau, z = 0)$ of the orbifold model $\tilde{\mathcal{C}}$ (we underline the classes that restrict to $\mathbb{M}_{24}$ conjugacy classes):

Co_0-class	$\underline{\text{1A}}$	$\underline{\text{2B}}$	2C	$\underline{\text{3B}}$	3C	4B	$\underline{\text{4E}}$	4F	$\underline{\text{5B}}$	5C	6G	6H	6I	$\underline{\text{6K}}$	6L	6M	$\underline{\text{7B}}$
dim fix	24	16	8	12	6	8	10	6	8	4	6	6	6	8	4	4	6
$\mathrm{Tr}_{\mathbf{24}}(g)$	24	8	−8	6	−3	8	4	−4	4	−1	−4	4	5	2	−2	−1	3
$\tilde{\phi}(\tau, 0)$	24	24	0	24	0	24	24	0	24	0	0	24	24	24	0	0	24

Co_0-class	8D	$\underline{\text{8G}}$	8H	9C	10F	10G	10H	$\underline{\text{11A}}$	12I	12L	12N	12O	$\underline{\text{14C}}$	$\underline{\text{15D}}$
dim fix	4	6	4	4	4	4	4	4	4	4	4	4	4	4
$\mathrm{Tr}_{\mathbf{24}}(g)$	4	2	−2	3	−2	2	3	2	2	1	−2	2	1	1
$\tilde{\phi}(\tau, 0)$	24	24	0	24	0	24	24	24	24	24	0	24	24	24

Note that the elliptic genus of the orbifold theory $\tilde{\mathcal{C}}$ is always 0 or 24, corresponding to a torus or a K3 sigma-model, respectively. Out of curiosity, we have also computed the putative elliptic genus $\tilde{\phi}(\tau, 0)$ for the 11 classes of symmetries g with $\mathrm{Tr}_{24}(g) = 0$ for which we do not expect the orbifold to make sense — the corresponding twining genus ϕ_g will typically have a non-trivial multiplier system, and hence the orbifold will not satisfy level-matching. Indeed, for almost none of these cases is $\tilde{\phi}(\tau, 0)$ equal to 0 or 24, thus signaling an inconsistency of the orbifold model:

Co_0-class	$\underline{\text{2D}}$	$\underline{\text{3D}}$	4D	$\underline{\text{4G}}$	$\underline{\text{4H}}$	6O	$\underline{\text{6P}}$	8C	8I	$\underline{\text{10J}}$	$\underline{\text{12P}}$
dim fix	12	8	4	8	6	6	4	4	4	4	4
$\mathrm{Tr}_{\mathbf{24}}(g)$	0	0	0	0	0	0	0	0	0	0	0
$\tilde{\phi}(\tau, 0)$	12	8	0	12	6	12	4	12	6	12	12

The only exception is the class 4D, which might define a consistent orbifold (a torus model). It follows that a K3 model $\mathcal{C}$ is the $\mathbb{Z}_n$-orbifold of a torus model if and only if it contains a symmetry g in one of the classes

$$\begin{aligned} &\text{2C, 3C, 4F, 5C, 6G, 6L, 6M, 8H, 10F, 12N,} \\ &\text{4B, 4D, 6H, 6I, 8C, 8D, 9C, 10G, 10H, 12I, 12L, 12O} \end{aligned} \tag{4.4}$$

of Co_0.[¶] Here we have also included (in the second line) those classes of elements $g \in \mathrm{Co}_0$ for which $\mathcal{C}/\langle g^i \rangle$ is a torus model, for some power $i > 1$. Our main observation is now that none of the Co_0 classes in (4.4) restricts to a class in $\mathbb{M}_{24}$, i.e.

All K3 models that are realised as $\mathbb{Z}_n$-orbifolds of torus models are exceptional. In particular, the quantum symmetry is not an element of $\mathbb{M}_{24}$.

One might ask whether the converse is also true, i.e. whether all exceptional models are cyclic torus orbifolds. This is not quite the case: for example, the classification theorem of section 3 predicts the existence of models with a symmetry group $G \cong GL_2(3)$ (the group of 2×2 invertible matrices on the field $\mathbb{F}_3$ with 3 elements). The group G contains no elements in the classes (4.4), so the model is not a cyclic torus orbifold; on the other hand, G contains elements in the class 8I of Co_0, which does not restrict to $\mathbb{M}_{24}$. A second counterexample is a family of models with a symmetry g in the class 6O of Co_0. A generic point of this family is not a cyclic torus model (although some special points are), since the full symmetry group is generated by g and contains no elements in (4.4). Both these counterexamples belong to case (i) of the general classification theorem. In fact, we can prove that

---

[¶] We should emphasise that for us the term 'orbifold' always refers to a conformal field theory (rather than a geometrical) construction. Although a non-linear sigma-model on a geometric orbifold $M/\mathbb{Z}_N$ always admits an interpretation as a CFT orbifold, the converse is not always true. In particular, there exist asymmetric orbifold constructions that do not have a direct geometric interpretation, see for example section 5.

The symmetry group G of a K3 sigma-model $\mathcal{C}$ contains a subgroup $3^{1+4} : \mathbb{Z}_2$ (cases (iii) and (iv) of the theorem) if and only if $\mathcal{C}$ is a $\mathbb{Z}_3$-orbifold of a torus model. Furthermore, $G = 5^{1+2} : \mathbb{Z}_4$ (case (ii)) if and only if $\mathcal{C}$ is a $\mathbb{Z}_5$-orbifold of a torus model.

The proof goes as follows. All subgroups of Co_0 of the form $3^{1+4} : \mathbb{Z}_2$ (respectively, $5^{1+2} : \mathbb{Z}_4$) contain an element in the class 3C (resp., 5C), and therefore the corresponding models are $\mathbb{Z}_3$ (resp., $\mathbb{Z}_5$) torus orbifolds. Conversely, consider a $\mathbb{Z}_3$-orbifold of a torus model. Its symmetry group G contains the quantum symmetry g in class 3C of Co_0. (It must contain a symmetry generator of order three whose orbifold leads to a torus, and 3C is then the only possibility.) The sublattice $\Lambda^{\langle g\rangle} \subset \Lambda$ fixed by g is the $\mathcal{S}$-lattice $2^{27}3^{36}$ [31]. From the classification theorem, we know that G is the stabiliser of a sublattice $\Lambda^G \subset \Lambda$ of rank at least 4. Since $\Lambda^G \subseteq \Lambda^{\langle g\rangle}$, G contains as a subgroup the stabiliser of $\Lambda^{\langle g\rangle}$, namely $3^{1+4} : \mathbb{Z}_2$.

Analogously, a $\mathbb{Z}_5$ torus orbifold always has a symmetry in class 5C, whose fixed sublattice $\Lambda^{\langle g\rangle}$ is the $\mathcal{S}$-lattice $2^5 3^{10}$ [31]. Since $\Lambda^{\langle g\rangle}$ has rank 4 and is primitive, $\Lambda^G = \Lambda^{\langle g\rangle}$ and the symmetry group G must be the stabiliser $5^{1+2} : \mathbb{Z}_4$ of $\Lambda^{\langle g\rangle}$.

It was shown in [7] that the Gepner model $(1)^6$ corresponds to the case (ii) of the classification theorem. It thus follows from the above reasoning that it must indeed be equivalent to a $\mathbb{Z}_3$-orbifold of $\mathbb{T}^4$, see also [27]. (We shall also study these orbifolds more systematically in section 6.) In the next section, we will provide an explicit construction of a $\mathbb{Z}_5$-orbifold of a torus model and show that its symmetry group is $5^{1+2} : \mathbb{Z}_4$, as predicted by the above analysis.

5 A K3 model with symmetry group $5^{1+2} : \mathbb{Z}_4$

In this section we will construct a supersymmetric sigma-model on $\mathbb{T}^4$ with a symmetry g of order 5 commuting with an $\mathcal{N} = (4, 4)$ superconformal algebra and acting asymmetrically on the left- and on the right-moving sector. The orbifold of this model by g will turn out to be a well-defined SCFT with $\mathcal{N} = (4, 4)$ (in particular, the level matching condition is satisfied) that can be interpreted as a non-linear sigma-model on K3. We will argue that the group of symmetries of this model is $G = 5^{1+2}.\mathbb{Z}_4$, one of the exceptional groups considered in the general classification theorem.

5.1 The torus model

Let us consider a supersymmetric sigma-model on the torus $\mathbb{T}^4$. Geometrically, we can characterise the theory in terms of a metric and a Kalb-Ramond field, but it is actually more convenient to describe it simply as a conformal field theory that is generated by the following fields: four left-moving $u(1)$ currents $\partial X^a(z)$, $a = 1, \ldots, 4$, four free fermions $\psi^a(z)$, $a = 1, \ldots, 4$, their right-moving analogues $\bar{\partial} X^a(\bar{z}), \tilde{\psi}^a(\bar{z})$, as well as some winding-momentum fields $V_\lambda(z, \bar{z})$ that are associated to vectors λ in an even unimodular lattice $\Gamma^{4,4}$ of signature $(4,4)$. The mode expansions of the left-moving fields are

$$\partial X^a(z) = \sum_{n\in\mathbb{Z}} \alpha_n z^{-n-1} , \qquad \psi^a = \sum_{n\in\mathbb{Z}+\nu} \psi_n z^{-n-\frac{1}{2}} , \tag{5.1}$$

where $\nu = 0, 1/2$ in the R- and NS-sector, respectively. Furthermore, we have the usual commutation relations

$$[\alpha^a_m, \alpha^b_n] = m\, \delta^{ab}\, \delta_{m,-n} \qquad\qquad \{\psi^a_m, \psi^b_n\} = \delta^{ab}\, \delta_{m,-n} . \tag{5.2}$$

Analogous statements also hold for the right-moving modes $\tilde{\alpha}_n$ and $\tilde{\psi}_n$. The vectors $\lambda \equiv (\lambda_L, \lambda_R) \in \Gamma^{4,4}$ describe the eigenvalues of the corresponding state with respect to the left- and right-moving zero modes α^a_0 and $\tilde{\alpha}^a_0$, respectively. In these conventions the inner product on $\Gamma^{4,4}$ is given as

$$(\lambda, \lambda') = \lambda_L \cdot \lambda'_L - \lambda_R \cdot \lambda'_R . \tag{5.3}$$

5.1.1 Continuous and discrete symmetries

Any torus model contains an $\hat{\mathfrak{su}}(2)_1 \oplus \hat{\mathfrak{su}}(2)_1 \oplus \hat{\mathfrak{u}}(1)^4$ current algebra, both on the left and on the right. Here, the $\hat{\mathfrak{u}}(1)^4$ currents are the ∂X^a, $a = 1, \ldots, 4$, while $\hat{\mathfrak{su}}(2)_1 \oplus \hat{\mathfrak{su}}(2)_1 = \hat{\mathfrak{so}}(4)_1$ is generated by the fermionic bilinears

$$a^3 := \bar{\psi}^{(1)}\psi^{(1)} + \bar{\psi}^{(2)}\psi^{(2)} \qquad a^+ := \bar{\psi}^{(1)}\bar{\psi}^{(2)} \qquad a^- := -\psi^{(1)}\psi^{(2)} , \tag{5.4}$$

$$\hat{a}^3 := \bar{\psi}^{(1)}\psi^{(1)} - \bar{\psi}^{(2)}\psi^{(2)} \qquad \hat{a}^+ := \bar{\psi}^{(1)}\psi^{(2)} \qquad \hat{a}^- := -\psi^{(1)}\bar{\psi}^{(2)} , \tag{5.5}$$

where

$$\psi^{(1)} = \tfrac{1}{\sqrt{2}}(\psi^1 + i\psi^2) \qquad\qquad \psi^{(2)} = \tfrac{1}{\sqrt{2}}(\psi^3 + i\psi^4) \tag{5.6}$$

$$\bar{\psi}^{(1)} = \tfrac{1}{\sqrt{2}}(\psi^1 - i\psi^2) \qquad\qquad \bar{\psi}^{(2)} = \tfrac{1}{\sqrt{2}}(\psi^3 - i\psi^4) . \tag{5.7}$$

At special points in the moduli space, where the $\Gamma^{4,4}$ lattice contains vectors of the form $(\lambda_L, 0)$ with $\lambda_L^2 = 2$, the bosonic $\mathfrak{u}(1)^4$ algebra is enhanced to some non-abelian algebra $\mathfrak{g}$ of rank 4. There are generically 16 (left-moving) supercharges; they form four $(\mathbf{2}, \mathbf{2})$ representations of the $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ zero mode algebra from (5.4) and (5.5). Altogether, the chiral algebra at generic points is a *large* $\mathcal{N} = 4$ superconformal algebra.

We want to construct a model with a symmetry g of order 5, acting non-trivially on the fermionic fields, and commuting with the *small* $\mathcal{N} = 4$ subalgebras both on the left and on the right. A small $\mathcal{N} = 4$ algebra contains an $\hat{\mathfrak{su}}(2)_1$ current algebra and four supercharges in two doublets of $\mathfrak{su}(2)$. The symmetry g acts by an $\mathrm{O}(4, \mathbb{R})$ rotation on the left-moving fermions ψ^a, preserving the anti-commutation relations (5.2). Without loss of generality, we may assume that $\psi^{(1)}$ and $\bar{\psi}^{(1)}$ are eigenvectors of g with eigenvalues ζ and ζ^{-1}, where ζ is a primitive fifth root of unity

$$\zeta^5 = 1 \ , \tag{5.8}$$

and that the $\hat{\mathfrak{su}}(2)_1$ algebra preserved by g is (5.4). This implies that g acts on all the fermionic fields by

$$\psi^{(1)} \mapsto \zeta\, \psi^{(1)} \ , \quad \bar{\psi}^{(1)} \mapsto \zeta^{-1}\bar{\psi}^{(1)} \ , \quad \psi^{(2)} \mapsto \zeta^{-1}\psi^{(2)} \ , \quad \bar{\psi}^{(2)} \mapsto \zeta\, \bar{\psi}^{(2)} \ . \tag{5.9}$$

Note that the action of g on the fermionic fields can be described by $e^{\frac{2\pi i k}{5}\hat{a}^3_0}$ for some $k = 1, \ldots, 4$, where $\hat{a}^3$ is the current in the algebra (5.5). The four g-invariant supercharges can then be taken to be

$$\sqrt{2}\sum_{i=1}^{2} J^{(i)}\bar{\psi}^{(i)} \ , \quad \sqrt{2}\sum_{i=1}^{2} \bar{J}^{(i)}\psi^{(i)} \ , \quad \sqrt{2}(\bar{J}^{(1)}\bar{\psi}^{(2)} - \bar{J}^{(2)}\bar{\psi}^{(1)}) \ , \quad \sqrt{2}(J^{(1)}\psi^{(2)} - J^{(2)}\psi^{(1)}) \ , \tag{5.10}$$

where $J^{(1)}, \bar{J}^{(1)}, J^{(2)}, \bar{J}^{(2)}$ are suitable (complex) linear combinations of the left-moving currents ∂X^a, $a = 1, \ldots, 4$. In order to preserve the four supercharges, g must act with the same eigenvalues on the bosonic currents

$$J^{(1)} \mapsto \zeta\, J^{(1)} \ , \quad \bar{J}^{(1)} \mapsto \zeta^{-1}\bar{J}^{(1)} \ , \quad J^{(2)} \mapsto \zeta^{-1}J^{(2)} \ , \quad \bar{J}^{(2)} \mapsto \zeta\, \bar{J}^{(2)} \ . \tag{5.11}$$

A similar reasoning applies to the right-moving algebra with respect to an eigenvalue $\tilde{\zeta}$, with $\tilde{\zeta}^5 = 1$. For the symmetries with a geometric interpretation, the action on the left- and right-moving bosonic currents is induced by an $\mathrm{O}(4, \mathbb{R})$-transformation on the scalar fields X^a, $a, = 1, \ldots, 4$, representing the coordinates on the torus; then ζ and $\tilde{\zeta}$ are necessarily equal. In our treatment, we want to allow for the more general case where $\zeta \neq \tilde{\zeta}$.

The action of g on J^a and $\tilde{J}^a$ induces an $\mathrm{O}(4, 4, \mathbb{R})$-transformation on the lattice $\Gamma^{4,4}$. The transformation g is a symmetry of the model if and only if it induces an automorphism on $\Gamma^{4,4}$. In particular, it must act by an invertible integral matrix on any lattice basis. The requirement that the trace of this matrix (and of any power of it) must be integral, leads to the condition that

$$2(\zeta^i + \zeta^{-i} + \tilde{\zeta}^i + \tilde{\zeta}^{-i}) \in \mathbb{Z} \ , \tag{5.12}$$

for all $i \in \mathbb{Z}$. For g of order 5, this condition is satisfied by

$$\zeta = e^{\frac{2\pi i}{5}} \qquad \text{and} \qquad \tilde{\zeta} = e^{\frac{4\pi i}{5}} \ , \tag{5.13}$$

and this solution is essentially unique (up to taking powers of it or exchanging the definition of ζ and ζ^{-1}). Eq. (5.13) shows that a supersymmetry preserving symmetry of order 5 is necessarily left-right asymmetric, and hence does not have a geometric interpretation.

It is now clear how to construct a torus model with the symmetries (5.9) and (5.11). First of all, we need an automorphism g of $\Gamma^{4,4}$ of order five. Such an automorphism must have eigenvalues $\zeta, \zeta^2, \zeta^3, \zeta^4$, each corresponding to two independent eigenvectors $v^{(1)}_{\zeta^i}, v^{(2)}_{\zeta^i}$, $i = 1, \ldots, 4$, in $\Gamma^{4,4} \otimes \mathbb{C}$. Given the discussion above, see in particular (5.13), we now require that the vectors

$$v^{(1)}_{\zeta^1} , \quad v^{(2)}_{\zeta^1} , \quad v^{(1)}_{\zeta^4} , \quad v^{(2)}_{\zeta^4} \tag{5.14}$$

span a positive-definite subspace of $\Gamma^{4,4} \otimes \mathbb{C}$ (i.e. correspond to the left-movers), while the vectors

$$v^{(1)}_{\zeta^2} , \quad v^{(2)}_{\zeta^2} , \quad v^{(1)}_{\zeta^3} , \quad v^{(2)}_{\zeta^3} \tag{5.15}$$

span a negative-definite subspace of $\Gamma^{4,4} \otimes \mathbb{C}$ (i.e. correspond to the right-movers).

An automorphism g with the properties above can be explicitly constructed as follows. Let us consider the real vector space with basis vectors $x_1, \ldots, x_4$, and $y_1, \ldots, y_4$, and define a linear map g of order 5 by

$$g(x_i) = x_{i+1} , \qquad g(y_i) = y_{i+1} , \qquad\qquad i = 1, \ldots, 3 , \tag{5.16}$$

and

$$g(x_4) = -(x_1 + x_2 + x_3 + x_4) , \qquad g(y_4) = -(y_1 + y_2 + y_3 + y_4) . \tag{5.17}$$

A g-invariant bilinear form on the space is uniquely determined by the conditions

$$(x_i, x_i) = 0 = (y_i, y_i) , \qquad i = 1, \ldots, 4 \tag{5.18}$$

and

$$(x_1, x_2) = 1 , \quad (x_1, x_3) = (x_1, x_4) = -1 , \qquad (y_1, y_2) = 1 , \quad (y_1, y_3) = (y_1, y_4) = -1 , \tag{5.19}$$

as well as

$$(x_1, y_1) = 1 , \quad \text{and} \quad (x_i, y_1) = 0 , \quad (i = 2, 3, 4) . \tag{5.20}$$

The lattice spanned by these basis vectors is an indefinite even unimodular lattice of rank 8 and thus necessarily isomorphic to $\Gamma^{4,4}$. The g-eigenvectors can be easily constructed in terms of the basis vectors and one can verify that the eigenspaces have the correct signature.

This torus model has an additional $\mathbb{Z}_4$ symmetry group that preserves the superconformal algebra and normalises the group generated by g. The generator h of this group acts by

$$h(x_i) := g^{1-i}(x_1 + x_4 + 2y_1 + y_2 + y_3 + y_4) , \tag{5.21}$$

$$h(y_i) := g^{1-i}(-2x_1 - x_2 - x_3 - x_4 - y_1 - y_3 - y_4) , \qquad i = 1, \ldots, 4 , \tag{5.22}$$

on the lattice vectors. The g-eigenvectors $v^{(a)}_{\zeta^i}$, $a = 1, 2$, $i = 1, \ldots, 4$ can be defined as

$$v^{(1)}_{\zeta^i} := \sum_{j=0}^{4} \zeta^{-ij} g^j(x_1 + h(x_1)) , \qquad v^{(2)}_{\zeta^i} := \sum_{j=0}^{4} \zeta^{-ij} g^j(x_1 - h(x_1)) , \tag{5.23}$$

so that

$$h(v^{(1)}_{\zeta^i}) = -v^{(2)}_{\zeta^{-i}} , \qquad h(v^{(2)}_{\zeta^i}) = v^{(1)}_{\zeta^{-i}} . \tag{5.24}$$

Correspondingly, the action of h on the left-moving fields is

$$\psi^{(1)} \mapsto -\psi^{(2)}\ , \quad \psi^{(2)} \mapsto \psi^{(1)}\ , \quad \bar{\psi}^{(1)} \mapsto -\bar{\psi}^{(2)}\ , \quad \bar{\psi}^{(2)} \mapsto \bar{\psi}^{(1)}\ , \tag{5.25}$$

$$J^{(1)} \mapsto -J^{(2)}\ , \quad J^{(2)} \mapsto J^{(1)}\ , \quad \bar{J}^{(1)} \mapsto -\bar{J}^{(2)}\ , \quad \bar{J}^{(2)} \mapsto \bar{J}^{(1)}\ , \tag{5.26}$$

and the action on the right-moving fields is analogous. It is immediate to verify that the generators of the superconformal algebra are invariant under this transformation.

5.2 The orbifold theory

Next we want to consider the orbifold of this torus theory by the group $\mathbb{Z}_5$ that is generated by g.

5.2.1 The elliptic genus

The elliptic genus of the orbifold theory can be computed by summing over the $\mathrm{SL}(2,\mathbb{Z})$ images of the untwisted sector contribution, which in turn is given by

$$\phi^U(\tau,z) = \frac{1}{5}\sum_{k=0}^{4} \phi_{1,g^k}(\tau,z)\ , \tag{5.27}$$

where

$$\phi_{1,g^k}(\tau,z) = \mathrm{Tr}_{\mathrm{RR}}(g^k\, q^{L_0-\frac{c}{24}}\bar{q}^{\tilde{L}_0-\frac{\tilde{c}}{24}} y^{2J_0}(-1)^{F+\tilde{F}})\ . \tag{5.28}$$

The $k=0$ contribution, i.e. the elliptic genus of the original torus theory, is zero. Each g^k-contribution, for $k=1,\ldots,4$, is the product of a factor coming from the ground states, one from the oscillators and one from the momenta

$$\phi_{1,g^k}(\tau,z) = \phi^{\mathrm{gd}}_{1,g^k}(\tau,z)\,\phi^{\mathrm{osc}}_{1,g^k}(\tau,z)\,\phi^{\mathrm{mom}}_{1,g^k}(\tau,z)\ . \tag{5.29}$$

These contributions are, respectively,

$$\phi^{\mathrm{gd}}_{1,g^k}(\tau,z) = y^{-1}(1-\zeta^k y)(1-\zeta^{-k}y)(1-\zeta^{2k})(1-\zeta^{-2k}) = 2y^{-1}+2y+1\ , \tag{5.30}$$

$$\phi^{\mathrm{osc}}_{1,g^k}(\tau,z) = \prod_{n=1}^{\infty}\frac{(1-\zeta^k y q^n)(1-\zeta^{-k}yq^n)(1-\zeta^k y^{-1}q^n)(1-\zeta^{-k}y^{-1}q^n)}{(1-\zeta^k q^n)^2(1-\zeta^{-k}q^n)^2}\ , \tag{5.31}$$

and

$$\phi^{\mathrm{mom}}_{1,g^k}(\tau,z) = 1\ , \tag{5.32}$$

where the last result follows because the only g-invariant state of the form (k_L,k_R) is the vacuum $(0,0)$. Thus we have

$$\phi_{1,g^k}(\tau,z) = 5\,\frac{\vartheta_1(\tau,z+\frac{k}{5})\,\vartheta_1(\tau,z-\frac{k}{5})}{\vartheta_1(\tau,\frac{k}{5})\,\vartheta_1(\tau,-\frac{k}{5})}\ , \tag{5.33}$$

where

$$\vartheta_1(\tau,z) = -iq^{1/8}y^{-\frac{1}{2}}(y-1)\prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^n)\ , \tag{5.34}$$

is the first Jacobi theta function. Modular transformations of $\phi_{1,g^k}(\tau,z)$ reproduce the twining genera in the twisted sector

$$\phi_{g^l,g^k}(\tau,z) = \mathrm{Tr}_{\mathcal{H}^{(l)}}\left(g^k\, q^{L_0-\frac{c}{24}}\,\bar{q}^{\tilde{L}_0-\frac{\tilde{c}}{24}}y^{2J_0}(-1)^{F+\tilde{F}}\right)\ , \tag{5.35}$$

and using the modular properties of the theta function we obtain

$$\phi_{g^l,g^k}(\tau,z) = 5\,\frac{\vartheta_1(\tau,z+\frac{k}{5}+\frac{l\tau}{5})\,\vartheta_1(\tau,z-\frac{k}{5}-\frac{l\tau}{5})}{\vartheta_1(\tau,\frac{k}{5}+\frac{l\tau}{5})\,\vartheta_1(\tau,-\frac{k}{5}-\frac{l\tau}{5})}\ , \tag{5.36}$$

for $k,l\in\mathbb{Z}/5\mathbb{Z}$, $(k,l)\neq(0,0)$. The elliptic genus of the full orbifold theory is then

$$\phi_{\rm orb}(\tau,z) = \frac{1}{5}\sum_{k,l\in\mathbb{Z}/5\mathbb{Z}}\phi_{g^l,g^k}(\tau,z) = \sum_{\substack{k,l\in\mathbb{Z}/5\mathbb{Z}\\(k,l)\neq(0,0)}}\frac{\vartheta_1(\tau,z+\frac{k}{5}+\frac{l\tau}{5})\,\vartheta_1(\tau,z-\frac{k}{5}-\frac{l\tau}{5})}{\vartheta_1(\tau,\frac{k}{5}+\frac{l\tau}{5})\,\vartheta_1(\tau,-\frac{k}{5}-\frac{l\tau}{5})}\ . \tag{5.37}$$

Since $\phi_{g^k,g^l}(\tau,0)=5$ for all $(k,l)\neq(0,0)$, we have

$$\phi_{\rm orb}(\tau,0) = \frac{1}{5}\sum_{\substack{k,l\in\mathbb{Z}/5\mathbb{Z}\\(k,l)\neq(0,0)}} 5 = 24\ , \tag{5.38}$$

which shows that the orbifold theory is a non-linear sigma-model on K3. In particular, the untwisted sector has 4 RR ground states, while each of the four twisted sectors contains 5 RR ground states. For the following it will be important to understand the structure of the various twisted sectors in detail.

5.2.2 The twisted sectors

In the g^k-twisted sector, let us consider a basis of g-eigenvectors for the currents and fermionic fields. For a given eigenvalue ζ^i, $i\in\mathbb{Z}/5\mathbb{Z}$, of g^k, the corresponding currents $J^{i,a}$ and fermionic fields $\psi^{i,b}$ (where a,b labels distinct eigenvectors with the same eigenvalue) have a mode expansion

$$J^{i,a}(z) = \sum_{n\in\frac{i}{5}+\mathbb{Z}}\alpha^{i,a}_n z^{-n-1}\ ,\qquad \psi^{i,a}(z) = \sum_{r\in\frac{i}{5}+\nu+\mathbb{Z}}\psi^{i,a}_r z^{-r-1/2}\ , \tag{5.39}$$

where $\nu=1/2$ in the NS- and $\nu=0$ in the R-sector. The ground states of the g^k-twisted sector are characterised by the conditions

$$\alpha^{i,a}_n|m,k\rangle = \tilde{\alpha}^{i,a}_n|m,k\rangle = 0\ ,\qquad \forall\ n>0,\ i,\ a\ , \tag{5.40}$$

$$\psi^{i,b}_r|m,k\rangle = \tilde{\psi}^{i,b}_r|m,k\rangle = 0\ ,\qquad \forall\ r>0,\ i,\ b\ , \tag{5.41}$$

where $|m,k\rangle$ denotes the $m^{\rm th}$ ground state in the g^k-twisted sector. Note that since none of the currents are g-invariant, there are no current zero modes in the g^k-twisted sector, and similarly for the fermions. For a given k, the states $|m,k\rangle$ have then all the same conformal dimension, which can be calculated using the commutation relation $[L_{-1},L_1]=2L_0$ or read off from the leading term of the modular transform of the twisted character (5.35). In the g^k-twisted NS-NS-sector the ground states have conformal dimension

$$\text{NS-NS } g^k\text{-twisted:} \qquad h = \frac{k}{5} \qquad \text{and} \qquad \tilde{h} = \frac{2k}{5} \ , \tag{5.42}$$

while in the RR-sector we have instead

$$\text{R-R } g^k\text{-twisted:} \qquad h = \tilde{h} = \frac{1}{4} \ . \tag{5.43}$$

In particular, level matching is satisfied, and thus the asymmetric orbifold is consistent [32]. The full g^k-twisted sector is then obtained by acting with the negative modes of the currents and the fermionic fields on the ground states $|m,k\rangle$.

Let us have a closer look at the ground states of the g^k twisted sector; for concreteness we shall restrict ourselves to the case $k=1$, but the modifications for general k are minor (see below). The vertex operators $V_\lambda(z,\bar{z})$ associated to $\lambda \in \Gamma^{4,4}$, act on the ground states $|m,1\rangle$ by

$$\lim_{z\to 0} V_\lambda(z,\bar{z})|m,1\rangle = e_\lambda |m,1\rangle \ , \tag{5.44}$$

where e_λ are operators commuting with all current and fermionic oscillators and satisfying

$$e_\lambda \, e_\mu = \epsilon(\lambda,\mu) \, e_{\lambda+\mu} \ , \tag{5.45}$$

for some fifth root of unity $\epsilon(\lambda,\mu)$. The vertex operators V_λ and V_μ must be local relative to one another, and this is the case provided that (see the appendix)

$$\frac{\epsilon(\lambda,\mu)}{\epsilon(\mu,\lambda)} = C(\lambda,\mu) \ , \tag{5.46}$$

where

$$C(\lambda,\mu) = \prod_{i=1}^{4} (\zeta^i)^{(g^i(\lambda),\mu)} = \zeta^{(P_g(\lambda),\mu)} \qquad \text{with} \qquad P_g(\lambda) = \sum_{i=1}^{4} i g^i(\lambda) \ . \tag{5.47}$$

The factor $C(\lambda,\mu)$ has the properties

$$C(\lambda,\mu_1+\mu_2) = C(\lambda,\mu_1)\,C(\lambda,\mu_2) \ , \qquad C(\lambda_1+\lambda_2,\mu) = C(\lambda_1,\mu)\,C(\lambda_2,\mu) \ , \tag{5.48}$$

$$C(\lambda,\mu) = C(\mu,\lambda)^{-1} \ , \tag{5.49}$$

$$C(\lambda,\mu) = C(g(\lambda),g(\mu)) \ . \tag{5.50}$$

Because of (5.48), $C(\lambda,0) = C(0,\lambda) = 1$ for all $\lambda \in \Gamma^{4,4}$, and we can set

$$e_0 = 1 \ , \tag{5.51}$$

so that $\epsilon(0,\lambda) = 1 = \epsilon(\lambda,0)$. More generally, for the vectors λ in the sublattice

$$R := \{\lambda \in \Gamma^{4,4} \mid C(\lambda,\mu) = 1 \ , \text{ for all } \mu \in \Gamma^{4,4}\} \subset \Gamma^{4,4} \ , \tag{5.52}$$

we have $C(\lambda+\mu_1,\mu_2) = C(\mu_1,\mu_2)$, for all $\mu_1,\mu_2 \in \Gamma^{4,4}$, so that we can set

$$e_{\mu+\lambda} = e_\mu \ , \qquad \forall \lambda \in R \ , \quad \mu \in \Gamma^{4,4} \ . \tag{5.53}$$

Thus, we only need to describe the operators corresponding to representatives of the group $\Gamma^{4,4}/R$. The vectors $\lambda \in R$ are characterised by

$$(P_g(\lambda), \mu) \equiv 0 \mod 5 \ , \qquad \text{for all } \mu \in \Gamma^{4,4} \ , \tag{5.54}$$

and since $\Gamma^{4,4}$ is self-dual this condition is equivalent to

$$P_g(\lambda) \in 5\,\Gamma^{4,4} \ . \tag{5.55}$$

Since g has no invariant subspace, we have the identity

$$1 + g + g^2 + g^3 + g^4 = 0 \tag{5.56}$$

that implies (see (5.47))

$$P_g \circ (1-g) = (1-g) \circ P_g = -5 \cdot \mathbf{1} \ . \tag{5.57}$$

Thus, $\lambda \in R$ if and only if

$$P_g(\lambda) = P_g \circ (1-g)(\tilde{\lambda}) \ , \tag{5.58}$$

for some $\tilde{\lambda} \in \Gamma^{4,4}$, and since P_g has trivial kernel (see (5.57)), we finally obtain

$$R = (1-g)\,\Gamma^{4,4} \ . \tag{5.59}$$

Since also $(1-g)$ has trivial kernel, R has rank 8 and $\Gamma^{4,4}/R$ is a finite group. Furthermore,

$$|\Gamma^{4,4}/R| = \det(1-g) = 25 \ , \tag{5.60}$$

and, since $5\,\Gamma^{4,4} \subset R$, the group $\Gamma^{4,4}/R$ has exponent 5. The only possibility is

$$\Gamma^{4,4}/R \cong \mathbb{Z}_5 \times \mathbb{Z}_5 \ . \tag{5.61}$$

Let $x, y \in \Gamma^{4,4}$ be representatives for the generators of $\Gamma^{4,4}/R$. By (5.49), we know that $C(x,x) = C(y,y) = 1$, so that $C(x,y) \neq 1$ (otherwise C would be trivial over the whole $\Gamma^{4,4}$), and we can choose x, y such that

$$C(x,y) = \zeta \ . \tag{5.62}$$

Thus, the ground states form a representation of the algebra of operators generated by e_x, e_y, satisfying

$$e_x^5 = 1 = e_y^5 \ , \qquad e_x e_y = \zeta e_y e_x \ . \tag{5.63}$$

The group generated by e_x and e_y is the extra-special group 5^{1+2}, and all its non-abelian irreducible representations[‖] are five dimensional.

In particular, for the representation on the g-twisted ground states, we can choose a basis of e_x-eigenvectors

$$|m;1\rangle \quad \text{with} \quad e_x|m;1\rangle = \zeta^m |m;1\rangle \ , \qquad m \in \mathbb{Z}/5\mathbb{Z} \ , \tag{5.64}$$

[‖] We call a representation non-abelian if the central element does not act trivially.

and define the action of the operators e_y by

$$e_y|m;1\rangle = |m+1;1\rangle \ . \tag{5.65}$$

For any vector $\lambda \in \Gamma^{4,4}$, there are unique $a, b \in \mathbb{Z}/5\mathbb{Z}$ such that $\lambda = ax + by + (1-g)(\mu)$ for some $\mu \in \Gamma^{4,4}$ and we define[**]

$$e_\lambda := e_x^a e_y^b \ . \tag{5.66}$$

Since $g(\lambda) = ax + by + (1-g)(g(\mu) - ax - by)$, by (5.53) we have

$$e_{g(\lambda)} = e_\lambda \ , \tag{5.67}$$

so that, with respect to the natural action $g(e_\lambda) := e_{g(\lambda)}$, the algebra is g-invariant. This is compatible with the fact that all ground states have the same left and right conformal weights h and $\tilde{h}$, so that the action of $g = e^{2\pi i(h-\tilde{h})}$ is proportional to the identity.

The construction of the g^k-twisted sector, for $k = 2, 3, 4$, is completely analogous to the g^1-twisted case, the only difference being that the root ζ in the definition of $C(\lambda, \mu)$ should be replaced by ζ^k. Thus, one can define operators $e_x^{(k)}, e_y^{(k)}$ on the g^k-twisted sector, for each $k = 1, \ldots, 4$, satisfying

$$(e_x^{(k)})^5 = 1 = (e_y^{(k)})^5 \ , \qquad e_x^{(k)} e_y^{(k)} = \zeta^k e_y^{(k)} e_x^{(k)} \ . \tag{5.68}$$

The action of these operators on the analogous basis $|m;k\rangle$ with $m \in \mathbb{Z}/5\mathbb{Z}$ is then

$$e_x^{(k)}|m;k\rangle = \zeta^m |m;k\rangle \ , \quad e_y^{(k)}|m;k\rangle = |m+k;k\rangle \ . \tag{5.69}$$

5.2.3 Spectrum and symmetries

The spectrum of the actual orbifold theory is finally obtained from the above twisted sectors by projecting onto the g-invariant states; technically, this is equivalent to restricting to the states for which the difference of the left- and right- conformal dimensions is integer, $h - \tilde{h} \in \mathbb{Z}$. In particular, the RR ground states (5.69) in each (twisted) sector have $h = \tilde{h} = 1/4$, so that they all survive the projection. Thus, the orbifold theory has 4 RR ground states in the untwisted sector (the spectral flow generators), forming a $(\mathbf{2}, \mathbf{2})$ representation of $\mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R$, and 5 RR ground states in each twisted sector, which are singlets of $\mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R$. In total there are therefore 24 RR ground states, as expected for a non-linear sigma-model on K3. (Obviously, we are here just reproducing what we already saw in (5.38).)

Next we want to define symmetry operators acting on the orbifold theory. First we can construct operators e_λ associated to $\lambda \in \Gamma^{4,4}$, that will form the extra special group 5^{1+2}. They are defined to act by $e_\lambda^{(k)}$ on the g^k-twisted sector. The action of the untwisted sector preserves the subspaces $\mathcal{H}_m^U$, $m \in \mathbb{Z}/5\mathbb{Z}$, of states with momentum of the form $\lambda = nx + my + (1-g)(\mu)$, for some $n \in \mathbb{Z}$ and $\mu \in \Gamma^{4,4}$. Let us denote by $T_{m;k}$ a generic vertex operator associated with a g^k-twisted state, $k = 1, \ldots, 4$, with e_x-eigenvalue ζ^m,

---

[**] The ordering of e_x and e_y in this definition is arbitrary; however, any other choice corresponds to a redefinition $\tilde{e}_\lambda = c(\lambda)e_\lambda$, for some fifth root of unity $c(\lambda)$, that does not affect the commutation relations $\tilde{e}_\lambda \tilde{e}_\mu = C(\lambda, \mu)\tilde{e}_\mu \tilde{e}_\lambda$.

$m \in \mathbb{Z}/5\mathbb{Z}$, and by $T_{m;0}$ a vertex operator associated with a state in $\mathcal{H}^U_m$. Consistency of the OPE implies the fusion rules

$$T_{m;k} \times T_{m';k'} \to T_{m+m';k+k'} \ . \tag{5.70}$$

These rules are preserved by the maps

$$T_{m;k} \mapsto e_\lambda \, T_{m;k} \, e_\lambda^{-1} \ , \qquad \lambda \in \Gamma^{4,4} \ , \tag{5.71}$$

which therefore define symmetries of the orbifold theory. As we have explained above, these symmetries form the extra-special group 5^{1+2}.

Finally, the symmetries (5.21), (5.22), (5.25) and (5.26) of the original torus theory induce a $\mathbb{Z}_4$-group of symmetries of the orbifold. Since $h^{-1}gh = g^{-1}$, the space of g-invariant states of the original torus theory is stabilised by h, so that h restricts to a well-defined transformation on the untwisted sector of the orbifold. Furthermore, h maps the g^k- to the g^{5-k}-twisted sector. Eqs. (5.21) and (5.22) can be written as

$$h(x_1) = 2x_1 + (1-g)(-x_1 - x_2 - x_3 + y_1 + y_2 + y_3 + y_4) \ , \tag{5.72}$$
$$h(y_1) = 2y_1 + (1-g)(-x_1 - x_2 - x_3 - x_4 - 2y_1 - y_2 - y_3 - y_4) \ . \tag{5.73}$$

It follows that the action of h on the operators $e^{(k)}_\lambda$, $k = 1, \ldots, 4$ must be

$$h e^{(k)}_x h^{-1} = e^{(5-k)}_{2x} \ , \qquad h e^{(k)}_y h^{-1} = e^{(5-k)}_{2y} \ , \tag{5.74}$$

and it is easy to verify that this transformation is compatible with (5.68). Correspondingly, the action on the twisted sector ground states is

$$h|m;k\rangle = |3m; 5-k\rangle \ , \tag{5.75}$$

and it is consistent with (5.70).

Thus the full symmetry group is the semi-direct product

$$G = 5^{1+2} : \mathbb{Z}_4 \ , \tag{5.76}$$

where the generator $h \in \mathbb{Z}_4$ maps the central element $\zeta \in 5^{1+2}$ to ζ^{-1}.

All of these symmetries act trivially on the superconformal algebra and on the spectral flow generators, and therefore define symmetries in the sense of the general classification theorem. Indeed, G agrees precisely with the group in case (ii) of the theorem. Thus our orbifold theory realises this possibility.

6 Models with symmetry group containing $3^{1+4} : \mathbb{Z}_2$

Most of the torus orbifold construction described in the previous section generalises to symmetries g of order different than 5. In particular, one can show explicitly that orbifolds of $\mathbb{T}^4$ models by a symmetry g of order 3 contain a group of symmetries $3^{1+4} : \mathbb{Z}_2$, so that they belong to one of the cases (iii) and (iv) of the theorem, as expected from the discussion in section 4.

We take the action of the symmetry g on the left-moving currents and fermionic fields to be of the form (5.9) and (5.11), where ζ is a now a third root of unity; analogous transformations hold for the right-moving fields with respect to a third root of unity $\tilde{\zeta}$. In this case, eq.(5.12) can be satisifed by

$$\zeta = \tilde{\zeta} = e^{\frac{2\pi i}{3}} , \tag{6.1}$$

so that the action is left-right symmetric and g admits an interpretation as a geometric $\mathrm{O}(4,\mathbb{R})$-rotation of order 3 of the torus $\mathbb{T}^4$. For example, the torus $\mathbb{R}^4/(A_2 \oplus A_2)$, where A_2 is the root lattice of the $su(3)$ Lie algebra, with vanishing Kalb-Ramond field, admits such an automorphism.

The orbifold by g contains 6 RR ground states in the untwisted sector. In the k^{th} twisted sector, $k = 1, 2$, the ground states form a representation of an algebra of operators $e^{(k)}_\lambda$, $\lambda \in \Gamma^{4,4}$, satisfying the commutation relations

$$e^{(k)}_\lambda e^{(k)}_\mu = C(\lambda,\mu)^k e^{(k)}_\mu e^{(k)}_\lambda , \tag{6.2}$$

where

$$C(\lambda,\mu) = \zeta^{(P_g(\lambda),\mu)} , \qquad P_g = g + 2g^2 . \tag{6.3}$$

As discussed in section 5.2.2, we can set

$$e^{(k)}_{\lambda+\mu} = e^{(k)}_\mu , \qquad \forall \lambda \in R, \mu \in \Gamma^{4,4} , \tag{6.4}$$

where

$$R = (1-g)\Gamma^{4,4} . \tag{6.5}$$

(Note that $\Gamma^{4,4}$ contains no g-invariant vectors). The main difference with the analysis of section 5.2.2 is that, in this case,

$$\Gamma^{4,4}/R \cong \mathbb{Z}_3^4 . \tag{6.6}$$

In particular, we can find vectors $x_1, x_2, y_1, y_2 \in \Gamma^{4,4}$ such that

$$C(x_i, y_j) = \zeta^{\delta_{ij}} , \qquad C(x_i, x_j) = C(y_i, y_j) = 1 . \tag{6.7}$$

The corresponding operators obey the relations

$$e^{(k)}_{x_i} e^{(k)}_{y_j} = \zeta^{k\delta_{ij}} e^{(k)}_{y_j} e^{(k)}_{x_i} , \quad e^{(k)}_{x_i} e^{(k)}_{x_j} = e^{(k)}_{x_j} e^{(k)}_{x_i} , \quad e^{(k)}_{y_i} e^{(k)}_{y_j} = e^{(k)}_{y_j} e^{(k)}_{y_i} , \tag{6.8}$$

as well as

$$(e^{(k)}_{x_i})^3 = 1 = (e^{(k)}_{y_i})^3 . \tag{6.9}$$

These operators generate the extra-special group 3^{1+4} of exponent 3, and the k^{th}-twisted ground states form a representation of this group. We can choose a basis $|m_1, m_2; k\rangle$, with $m_1, m_2 \in \mathbb{Z}/3\mathbb{Z}$, of simultaneous eigenvectors of $e^{(k)}_{x_1}$ and $e^{(k)}_{x_2}$, so that

$$e^{(k)}_{x_i} |m_1, m_2; k\rangle = \zeta^{m_i} |m_1, m_2; k\rangle , \qquad e^{(k)}_{y_i} |m_1, m_2; k\rangle = |m_1 + k\delta_{1i}, m_2 + k\delta_{2i}; k\rangle . \tag{6.10}$$

The resulting orbifold model has 9 RR ground states in each twisted sector, for a total of $6 + 9 + 9 = 24$ RR ground states, as expected for a K3 model. As in section 5.2.3, the group 3^{1+4} generated by $e_\lambda^{(k)}$ extends to a group of symmetries of the whole orbifold model. Furthermore, the $\mathbb{Z}_2$-symmetry that flips the signs of the coordinates in the original torus theory induces a symmetry h of the orbifold theory, which acts on the twisted sectors by

$$h\,|m_1, m_2; k\rangle = |-m_1, -m_2; k\rangle \ . \tag{6.11}$$

We conclude that the group G of symmetries of any torus orbifold $\mathbb{T}^4/\mathbb{Z}_3$ contains a subgroup $3^{1+4} : \mathbb{Z}_2$, and is therefore included in the cases (iii) or (iv) of the classification theorem. This obviously ties in nicely with the general discussion of section 4.

7 Conclusions

In this paper we have reviewed the current status of the EOT conjecture concerning a possible $\mathbb{M}_{24}$ symmetry appearing in the elliptic genus of K3. We have explained that, in some sense, the EOT conjecture has already been proven since twining genera, satisfying the appropriate modular and integrality properties, have been constructed for all conjugacy classes of $\mathbb{M}_{24}$. However, the analogue of the Monster conformal field theory that would 'explain' the underlying symmetry has not yet been found. In fact, no single K3 sigma-model will be able to achieve this since none of them possesses an automorphism group that contains $\mathbb{M}_{24}$.

Actually, the situation is yet further complicated by the fact that there are K3 sigma-models whose automorphism group is *not even a subgroup of* $\mathbb{M}_{24}$; on the other hand, the elliptic genus of K3 does not show any signs of exhibiting 'Moonshine' with respect to any larger group. As we have explained in this paper, most of the exceptional automorphism groups (i.e. automorphism groups that are not subgroups of $\mathbb{M}_{24}$) appear for K3s that are torus orbifolds. In fact, all cyclic torus orbifolds are necessarily exceptional in this sense, and (cyclic) torus orbifolds account for all incarnations of the cases (ii) – (iv) of the classification theorem of [7] (see section 4). We have checked these predictions by explicitly constructing an asymmetric $\mathbb{Z}_5$ orbifold that realises case (ii) of the theorem (see section 5), and a family of $\mathbb{Z}_3$ orbifolds realising cases (iii) and (iv) of the theorem (see section 6). Incidentally, these constructions also demonstrate that the exceptional cases (ii)-(iv) actually appear in the K3 moduli space — in the analysis of [7] this conclusion relied on some assumption about the regularity of K3 sigma-models.

The main open problem that remains to be understood is why precisely $\mathbb{M}_{24}$ is 'visible' in the elliptic genus of K3, rather than any smaller (or indeed bigger) group. Recently, we have constructed (some of) the twisted twining elliptic genera of K3 [33], i.e. the analogues of Simon Norton's generalised Moonshine functions [34]. We hope that they will help to shed further light on this question.

Acknowledgements

We thank Terry Gannon, Gerald Hoehn, and Michael Tuite for useful conversations. The research of MRG is supported in parts by the Swiss National Science Foundation.

A Commutation relations in the twisted sector

The vertex operators $V_\lambda(z,\bar z)$ in the g-twisted sector can be defined in terms of formal exponentials of current oscillators

$$E^{\pm}_\lambda(z,\bar z) := \exp\Big(\sum_{\substack{r\in\frac{1}{5}\mathbb{Z}\\ \pm r>0}} (\lambda_L\cdot\alpha)^{(r)}_r \frac{z^{-r}}{r}\Big) \exp\Big(\sum_{\substack{r\in\frac{1}{5}\mathbb{Z}\\ \pm r>0}} (\lambda_R\cdot\tilde\alpha)^{(r)}_r \frac{\bar z^{-r}}{r}\Big)\ , \tag{A.1}$$

where $(\lambda_L\cdot\alpha)^{(r)}_r$ and $(\lambda_R\cdot\tilde\alpha)^{(r)}_r$ are the r-modes of the currents

$$(\lambda_L\cdot\partial X)^{(r)} := \frac{1}{5}\sum_{i=0}^{4}\zeta^{5ir}\lambda_L\cdot g^i(\partial X) = \frac{1}{5}\sum_{i=0}^{4}\zeta^{5ir} g^{-i}(\lambda_L)\cdot\partial X\ , \tag{A.2}$$

$$(\lambda_R\cdot\bar\partial X)^{(r)} := \frac{1}{5}\sum_{i=0}^{4}\bar\zeta^{5ir}\lambda_R\cdot g^i(\bar\partial X) = \frac{1}{5}\sum_{i=0}^{4}\bar\zeta^{5ir} g^{-i}(\lambda_R)\cdot\bar\partial X\ . \tag{A.3}$$

Then we can define

$$V_\lambda(z,\bar z) := E^-_\lambda(z,\bar z)\, E^+_\lambda(z,\bar z)\, e_\lambda\ , \tag{A.4}$$

where the operators e_λ commute with all current oscillators and satisfy

$$e_\lambda\, e_\mu = \epsilon(\lambda,\mu)\, e_{\lambda+\mu}\ , \tag{A.5}$$

for some fifth root of unity $\epsilon(\lambda,\mu)$. The commutator factor

$$C(\lambda,\mu) := \frac{\epsilon(\lambda,\mu)}{\epsilon(\mu,\lambda)}\ , \tag{A.6}$$

can be determined by imposing the locality condition

$$V_\lambda(z_1,\bar z_1)\, V_\mu(z_2,\bar z_2) = V_\mu(z_2,\bar z_2)\, V_\lambda(z_1,\bar z_1)\ . \tag{A.7}$$

To do so, we note that the commutation relations between the operators $E^\pm_\lambda$ can be computed, as in [35], using the Campbell-Baker-Hausdorff formula

$$\begin{aligned} E^+_\lambda(z_1,\bar z_1)E^-_\mu(z_2,\bar z_2) &= E^-_\mu(z_2,\bar z_2)E^+_\lambda(z_1,\bar z_1)\\ &\qquad \prod_{i=0}^{4}[(1-\zeta^{-i}(\tfrac{z_1}{z_2})^{\frac{1}{5}})^{g^i(\lambda)_L\cdot\mu_L}(1-\bar\zeta^{-i}(\tfrac{\bar z_1}{\bar z_2})^{\frac{1}{5}})^{g^i(\lambda)_R\cdot\mu_R}]\ . \end{aligned} \tag{A.8}$$

Using (A.8) and $e_\lambda e_\mu = C(\lambda,\mu)e_\mu e_\lambda$, the locality condition is then equivalent to

$$C(\lambda,\mu)\prod_{i=0}^{4}\frac{(1-\zeta^{-i}(\frac{z_1}{z_2})^{\frac{1}{5}})^{g^i(\lambda)_L\cdot\mu_L}(1-\bar\zeta^{-i}\frac{\bar z_1^{1/5}}{\bar z_2^{1/5}})^{g^i(\lambda)_R\cdot\mu_R}}{(1-\zeta^{i}(\frac{z_2}{z_1})^{\frac{1}{5}})^{g^i(\lambda)_L\cdot\mu_L}(1-\bar\zeta^{i}(\frac{\bar z_2}{\bar z_1})^{\frac{1}{5}})^{g^i(\lambda)_R\cdot\mu_R}} = 1\ , \tag{A.9}$$

that is

$$C(\lambda,\mu)\Big(-\frac{z_1^{1/5}}{z_2^{1/5}}\Big)^{\sum_i g^i(\lambda)_L\cdot\mu_L}\Big(-\frac{\bar z_1^{1/5}}{\bar z_2^{1/5}}\Big)^{\sum_i g^i(\lambda)_R\cdot\mu_R}\prod_{i=0}^{4}[(\zeta^{-i})^{g^i(\lambda)_L\cdot\mu_L}(\bar\zeta^{-i})^{g^i(\lambda)_R\cdot\mu_R}] = 1\ . \tag{A.10}$$

Since $\Gamma^{4,4}$ has no g-invariant vector, we have the identities

$$\sum_{i=0}^{4} g^i(\lambda)_L = 0 = \sum_{i=0}^{4} g^i(\lambda)_R \ , \tag{A.11}$$

and hence finally obtain

$$C(\lambda,\mu) = \prod_{i=0}^{4} (\zeta^i)^{g^i(\lambda_L)\cdot\mu_L - g^i(\lambda_R)\cdot\mu_R} \ . \tag{A.12}$$

References

[1] T. Eguchi, H. Ooguri and Y. Tachikawa, *Notes on the K3 surface and the Mathieu group M_{24}*, Exper. Math. **20** (2011) 91 [`arXiv:1004.0956` `[hep-th]`].

[2] J.G. Thompson, *Some numerology between the Fischer-Griess Monster and the elliptic modular function*, Bull. Lond. Math. Soc. **11** (1979) 352.

[3] J.H. Conway and S. Norton, *Monstrous Moonshine*, Bull. Lond. Math. Soc. **11** (1979) 308.

[4] T. Gannon, *Moonshine beyond the Monster: The Bridge connecting Algebra, Modular Forms and Physics*, Cambridge University Press (2006).

[5] R. Borcherds, *Vertex algebras, Kac-Moody algebras and the Monster*, Proc. Nat. Acad. Sci. U.S.A. **83** (1986) 3068.

[6] I. Frenkel, J. Lepowski and A. Meurman, *Vertex Operator Algebras and the Monster*, Academic Press (1986).

[7] M.R. Gaberdiel, S. Hohenegger and R. Volpato, *Symmetries of K3 sigma models*, `arXiv:1106.4315` `[hep-th]`, to appear in Commun. Number Theory Phys.

[8] M.A. Walton, *Heterotic string on the simplest Calabi-Yau manifold and its orbifold limits*, Phys. Rev. D **37** (1988) 377.

[9] M.C.N. Cheng, *K3 Surfaces, N=4 dyons, and the Mathieu group M_{24}*, Commun. Number Theory Phys. **4** (2010) 623 [`arXiv:1005.5415` `[hep-th]`].

[10] M.C.N. Cheng and J.F.R. Duncan, *On Rademacher sums, the largest Mathieu group, and the holographic modularity of moonshine*, `arXiv:1110.3859` `[math.RT]`.

[11] M.C.N. Cheng and J.F.R. Duncan, *The largest Mathieu group and (mock) automorphic forms*, `arXiv:1201.4140` `[math.RT]`.

[12] M.C.N. Cheng, J.F.R. Duncan and J.A. Harvey, *Umbral moonshine*, `arXiv:1204.2779` `[math.RT]`.

[13] S. Govindarajan, *BKM Lie superalgebras from counting twisted CHL dyons*, JHEP **1105** (2011) 089 [`arXiv:1006.3472` `[hep-th]`].

[14] S. Govindarajan, *Brewing moonshine for Mathieu*, `arXiv:1012.5732 [math.NT]`.

[15] S. Govindarajan, *Unravelling Mathieu moonshine*, `arXiv:1106.5715 [hep-th]`.

[16] S. Hohenegger and S. Stieberger, *BPS saturated string amplitudes: K3 elliptic genus and Igusa cusp form*, Nucl. Phys. B **856** (2012) 413 [`arXiv:1108.0323 [hep-th]`].

[17] T. Kawai, Y. Yamada, and S.-K. Yang, *Elliptic genera and $\mathcal{N} = 2$ superconformal field theory*, Nucl. Phys. B **414** (1994) 191 [`arXiv:hep-th/9306096`].

[18] M. Eichler and D. Zagier, *The Theory of Jacobi Forms*, Birkhäuser (1985).

[19] S. Mukai, *Finite groups of automorphisms of $K3$ surfaces and the Mathieu group*, Invent. Math. **94** (1988) 183.

[20] S. Kondo, *Niemeier lattices, Mathieu groups and finite groups of symplectic automorphisms of K3 surfaces*, Duke Math. Journal **92** (1998) 593, appendix by S. Mukai.

[21] J.G. Thompson, *Finite groups and modular functions*, Bull. London Math. Soc. **11** (1979) 347.

[22] J.R. David, D.P. Jatkar and A. Sen, *Product representation of dyon partition function in CHL models*, JHEP **0606** (2006) 064 [`arXiv:hep-th/0602254`].

[23] M.R. Gaberdiel, S. Hohenegger and R. Volpato, *Mathieu moonshine in the elliptic genus of K3*, JHEP **1010** (2010) 062 [`arXiv:1008.3778 [hep-th]`].

[24] T. Eguchi and K. Hikami, *Note on twisted elliptic genus of K3 surface*, Phys. Lett. B **694** (2011) 446 [`arXiv:1008.4924 [hep-th]`].

[25] M.R. Gaberdiel, S. Hohenegger and R. Volpato, *Mathieu twining characters for K3*, JHEP **1009** (2010) 058 [`arXiv:1006.0221 [hep-th]`].

[26] T.J. Gannon, *Much ado about Mathieu*, to appear.

[27] M. Fluder, *Symmetries of non-linear sigma models on K3*, Master thesis, ETH Zürich, August 2011.

[28] M.P. Tuite, *Monstrous moonshine and the uniqueness of the moonshine module*, `arXiv:hep-th/9211069`.

[29] M.P. Tuite, *On the relationship between monstrous moonshine and the uniqueness of the moonshine module*, Commun. Math. Phys. **166** (1995) 495 [`arXiv:hep-th/9305057`].

[30] P.H. Ginsparg, *Applied conformal field theory*, `arXiv:hep-th/9108028`.

[31] R. T. Curtis, *On subgroups of $\cdot O$. II. Local structure*, J. Algebra **63** (1980) 413.

[32] K.S. Narain, M.H. Sarmadi and C. Vafa, *Asymmetric orbifolds*, Nucl. Phys. B **288** (1987) 551.

[33] M.R. Gaberdiel, D. Persson, H. Ronellenfitsch and R. Volpato, *in preparation.*

[34] S.P. Norton, *Generalised moonshine*, Proc. Symp. Pure Math. **47** (1987) 209.

[35] J. Lepowsky, *Calculus of twisted vertex operators*, Proc. Nat. Acad. Sci. U.S.A. **82** (1985) 8295.

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 45, Number 4, October 2008, Pages 675–679
S 0273-0979(08)01209-3
Article electronically published on June 25, 2008

Moonshine beyond the monster: The bridge connecting algebra, modular forms and physics, by Terry Gannon, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, Massachusetts, 2006, 492 pp., US$130.00, ISBN 978-0-521-83531-2

In about 1972 R. L. Griess and Fischer independently suggested the existence of a new sporadic simple group, which had a double cover of the "baby monster" discovered by Fischer as the centralizer of an involution. Calculations suggested that its smallest non-trivial complex representation probably had dimension 196883. "Moonshine" is the attempt to explain John McKay's extraordinary observation that this number is almost equal to the coefficient 196884 of the elliptic modular function

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + \cdots$$

giving the j-invariant of the elliptic curve $\mathbb{C}/\{1, \tau\}$ (where $q = e^{2\pi i\tau}$). At first, a common explanation was that there are many large numbers that turn up in mathematics, and a few of them will be almost equal just by coincidence.

Shortly afterwards, John Thompson [T] pointed out that the next coefficient 21493760 of the elliptic modular function is equal to the sum of the dimensions of the first 3 irreducible representations of the monster, and similarly all the other coefficients seems to be simple linear combinations of dimensions of irreducible representations. He conjectured that this might be because the monster acts on a graded representation V whose piece of degree n is the coefficient of q^n in $j(\tau) - 744$, and suggested looking at the traces of other elements of the monster on this representation. (The coefficient 744 of q^0 is a historical accident as adding an arbitrary constant to j still gives a function with similar properties. The most natural normalization is to set the constant term equal to 24, the number given by Rademacher's infinite series for coefficients of the j function.)

Conway and Norton [CN] carried out Thompson's suggestion and found that the traces of all elements of the monster seem to be given by the coefficients of Hauptmoduls (roughly, functions invariant under a genus 0 congruence subgroup of $SL_2(\mathbb{R})$). Atkin, Fong, and Smith [S] verified that the monster indeed has a representation with these properties by checking that the Hauptmoduls satisfied the necessary congruence and positivity conditions.

The monster simple group was finally constructed by R. L. Griess [G] in 1982. Norton had observed a few years before that the monster probably had a commutative (but non-associative) algebra structure on its 196883 dimensional representation, and it was also known how to write the 196884 dimensional representation as a sum of three irreducible representations of a centralizer of an involution. The monster could therefore be constructed by adjusting the finite number of parameters that an algebra structure depends upon so that the algebra has an extra automorphism. While the idea is simple, the details are horrendous to carry out: John Conway told me that when he heard Griess had constructed the monster, he assumed that Griess must have found another way to do it, because he could not believe that anyone would be patient enough to finish this calculation.

2000 *Mathematics Subject Classification.* Primary 17B69; Secondary 11F22, 20C34, 20D08.

©2008 American Mathematical Society
Reverts to public domain 28 years from publication

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Victor Kac [K] pointed out that the Lie group E_8 has a similar natural infinite dimensional representation, obtained by restricting an irreducible representation of the affine Lie algebra E_8, except that the traces of elements need not be Hauptmoduls but can be more general modular functions. More generally he showed that something similar happens for any highest weight representation of any affine Kac-Moody algebra when restricted to a suitable compact Lie group. Although the monster corresponds to a compact Lie group, no analogue of the corresponding affine Lie algebra is known for it.

Frenkel, Lepowsky, and Meurman [FLM] constructed an infinite dimensional module for the monster. This has the structure of a vertex algebra, and contains the Griess algebra inside it: one of the vertex algebra products is the Griess product. Some of the infinite dimensional representations of affine Lie algebras also carry similar vertex algebra structures, again showing that moonshine for the monster is very closely related to Kac's theory of affine Lie algebras.

The representation constructed by Frenkel, Lepowsky, and Meurman does indeed satisfy the moonshine conjectures; the proof of this uses a "generalized Kac-Moody algebra" acted on by the monster that is constructed using bosonic string theory. There is a small extension of Kac-Moody algebras called generalized Kac-Moody algebras, which unlike Kac-Moody algebras are allowed to have imaginary simple roots. Their theory is very similar: the basic theorems about them were proved by copying out Kac's proofs for the case of Kac-Moody algebras and making the obvious minor changes (though Kac pointed out to me that I introduced some completely original errors). The monster group does not exactly correspond to a Kac-Moody algebra; instead it acts as a group of diagram automorphisms of a generalized Kac-Moody algebra called the monster Lie algebra. The action of the monster group on the monster Lie algebra is similar to the "triality" action of the cyclic group of order 3 acting on the Lie algebra of the orthogonal group in 8 dimensions.

The proof of the moonshine conjecture does not directly show that the McKay-Thompson functions are genus 0. Instead it shows that they are completely replicable; this means that they satisfy identities similar to the product formula for the j-function. At first the fact that the completely replicable functions were Hauptmoduls was verified case by case, but more recently Cummins and Gannon [CG] found a more satisfactory uniform proof of this.

Some sporadic groups in the monster have natural graded representations that have vertex algebra structures only if they are reduced modulo some prime. For example, the baby monster has a representation on a graded space with character

$$q^{-1} + 4372q + 96256q^2 + \cdots$$

which cannot have a natural vertex algebra structure as the piece of dimension 4372 has no non-trivial product invariant under the baby monster. However Ryba [Ry] discovered that it probably does if reduced modulo 2, and in this case the whole space (probably) has a vertex algebra structure over the field with 2 elements invariant under the baby monster. Many, and possibly all, of the finite groups of Lie type also act naturally on vertex algebras over finite fields, which suggests the possibility that all finite simple groups act on suitable vertex algebras. However, for the sporadic groups not involved in the monster, no one has found any evidence of this, except possibly in the case of the Rudvalis group.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Conway and Norton [CNS] also found a particularly elegant presentation for the monster (or rather a closely related group) as a Coxeter group together with one simple extra relation. The fact that this is indeed a presentation for the monster was proved by A. A. Ivanov [I]. The relation between this presentation and the moonshine properties of the monster is unclear, though Miyamoto [M] has made some progress by showing that the generators of the presentation have a reasonably natural description as automorphisms of the monster vertex algebra; his construction uses the Niemeier lattice E_6^4, which makes one wonder if there are similar constructions for the other Niemeier lattices.

Another of McKay's strange observations about E_8 is that orbits of pairs of involutions of type 2A in the monster seem to correspond to the 9 nodes of the affine E_8 diagram. (There is a similar relation between the baby monster and the E_7 diagram, and between the sporadic group Fi_{24} and the E_6 diagram, so this is probably not a coincidence.) Recently Lam and Miyamoto [LM] found certain subalgebras of the monster vertex algebra corresponding to these nodes, which seems to go a long way towards explaining McKay's observation.

The proof of the moonshine conjectures depends on several coincidences. Even the existence of the monster seems to be a fluke in any of the known constructions: these all depend on long, strange calculations that just happen to work for no obvious reason, and would not have been done if the monster had not already been suspected to exist. Then the dimension of the Leech lattice just happens to be just 2 less than the critical dimension 26 of string theory, which is just what is needed for the no-ghost theorem to be used to construct the monster Lie algebra. The monster Lie algebra just happens to have a Weyl vector, which is extremely unusual for algebras constructed like this, and means that its simple roots can be described explicitly.

A major problem with studying the monster vertex algebra is that it is constructed by splicing together two quite different pieces. By analogy, the E_8 is also sometimes constructed as a sum of two pieces: a subalgebra D_8, and a half spin representation S of D_8. The Lie bracket on D_8 is already given, and the bracket of D_8 with S is given by the action of D_8 on S, and the bracket of two elements of S can be given by adjointness, using the fact that S has an invariant bilinear form. The resulting bracket on $D_8 \oplus S$ just happens to satisfy the Jacobi identity, so we get a Lie algebra E_8. The construction of the monster vertex algebra is formally very similar. In the case of E_8 there is a more natural construction which writes the E_8 Lie algebra as just one piece, rather than the sum of two different pieces. For this, one just takes the vertex algebra of (the double cover of) the E_8 lattice; the degree 1 piece of this is the E_8 Lie algebra. (It is not hard to eliminate vertex algebras from this construction and get a completely elementary construction of the E_8 Lie algebra as "one piece".) A major open problem is to find an analogous natural construction of the monster vertex algebra as just one piece, rather than as the sum of several unrelated smaller pieces. Duncan recently managed to do this for a similar superalgebra acted on by the double cover of Conway's simple group: he showed that the vertex superalgebra was isomorphic to the superalgebra of a certain 12 dimensional lattice. Later in [D], he used a similar idea to find a vertex algebra acted on by the Rudvalis groups: this is the first indication that it might be possible to extend moonshine to some of the 6 sporadic groups not involved in the monster.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

If a finite group acts on a vertex algebra, then the corresponding "orbifold" vertex algebra is easy to define: it is just the subalgebra fixed by the finite group. However it seems very hard to describe the representations of this subalgebra and the fusion between them, though there has been considerable progress by Dong, Li and Mason [DLM]. If this problem could be solved, it would quickly lead to a solution of the generalized moonshine conjectures, as it would be possible to construct analogues of the monster Lie algebra for all elements of the monster. One way of approaching orbifolds is via nets of von Neumann algebras, which seem to be closely related to vertex algebras: although there seems to be no natural functor going in either direction, whenever there is a construction of an interesting vertex algebra, there seems to be a similar construction of a net of von Neumann algebras, and vice versa. For example, Kawahigashi and Longo [KL] recently gave a construction of a net of von Neumann algebras with automorphism group the monster, by following the usual construction of the monster vertex algebra. For nets of von Neumann algebras, there is a good theory of orbifolds. Unfortunately it is not obvious how to use this to construct the Lie algebras associated to elements of the monster: the problem is that the elements of these Lie algebras correspond to unbounded operators, which are not so easy to get hold of using von Neumann algebras of bounded operators. So vertex algebras see the unbounded operators but cannot handle orbifolds well, while von Neumann algebras can handle orbifolds but cannot see the unbounded operators. Perhaps there is some way to combine vertex algebras and nets that can handle both orbifolds and unbounded operators, and therefore prove the generalized moonshine conjectures.

Hirzeburch [H] has suggested a "prize question" about the existence of a 24 dimensional manifold acted on by the monster, whose cohomology of twisted Dirac operators would give the monster vertex algebra. Hopkins and Mahowald [HM] showed the existence of a manifold with the topological properties suggested by Hirzebruch, though it does not seem to be known whether this manifold has a suitable action of the monster.

The most recent new idea about the monster is Witten's suggestion [W] of a relation between the monster and three dimensional gravity, which I am still too baffled by to say anything useful about.

At the moment we can prove almost anything we want about the monster and moonshine (or could with a bit more effort) but are really short of good explanations for what is going on. For example, there is still no expanation of why the monster exists that does not involve many pages of obscure calculations.

The book under review is a survey of some areas of mathematics related to moonshine. It is written in a lively and readable style, and the author explains what is going on instead of giving formal proofs. Much of the book is only indirectly related to the monster group: for example, there is a long chapter on conformal field theory, and an explanation of the relation between subfactors, knot polynomials, and quantum groups. It is an excellent introduction to this area for anyone who is looking for an informal survey.

One aspect of moonshine that is not discussed much in the present book is the relation between generalized Kac-Moody algebras of rank greater than 2 and automorphic forms of high rank. This is covered in a recent book [Ra] by U. Ray.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

REFERENCES

[CG] Cummins, C. J.; Gannon, T. Modular equations and the genus zero property of moonshine functions. Invent. Math. 129 (1997), no. 3, 413–443. MR1465329 (98k:11046)

[CN] Conway, J. H.; Norton, S. P. Monstrous moonshine. Bull. London Math. Soc. 11 (1979), no. 3, 308–339. MR554399 (81j:20028)

[CNS] Conway, John H.; Norton, Simon P.; Soicher, Leonard H. The Bimonster, the group Y_{555}, and the projective plane of order 3. Computers in algebra (Chicago, IL, 1985), 27–50, Lecture Notes in Pure and Appl. Math., 111, Dekker, New York, 1988. MR1060755 (92f:20018)

[DLM] Dong, Chongying; Li, Haisheng; Mason, Geoffrey. Modular-invariance of trace functions in orbifold theory and generalized Moonshine. Comm. Math. Phys. 214 (2000), no. 1, 1–56. MR1794264 (2001k:17043)

[D] Duncan, John F. Moonshine for Rudvalis's sporadic group I, arXiv:math/0609449

[FLM] Frenkel, I. B.; Lepowsky, J.; Meurman, A. A natural representation of the Fischer-Griess Monster with the modular function J as character. Proc. Nat. Acad. Sci. U.S.A. 81 (1984), no. 10, Phys. Sci., 3256–3260. MR747596 (85e:20018)

[G] Griess, Robert L., Jr. The friendly giant. Invent. Math. 69 (1982), no. 1, 1–102. MR671653 (84m:20024)

[H] Hirzebruch, Friedrich; Berger, Thomas; Jung, Rainer. Manifolds and modular forms. Aspects of Mathematics, E20. Friedr. Vieweg & Sohn, Braunschweig, 1992. xii+211 pp. ISBN: 3-528-06414-5 MR1189136 (94d:57001)

[HM] Mahowald, Mark; Hopkins, Mike. The structure of 24 dimensional manifolds having normal bundles which lift to BO[8]. Recent progress in homotopy theory (Baltimore, MD, 2000), 89–110, Contemp. Math., 293, Amer. Math. Soc., Providence, RI, 2002. MR1887530 (2003b:55007)

[I] Ivanov, A. A. A geometric characterization of the Monster. Groups, combinatorics & geometry (Durham, 1990), 46–62, London Math. Soc. Lecture Note Ser., 165, Cambridge Univ. Press, Cambridge, 1992. MR1200249 (94c:20033)

[K] Kac, V. G. An elucidation of: "Infinite-dimensional algebras, Dedekind's η-function, classical Möbius function and the very strange formula". $E_8^{(1)}$ and the cube root of the modular invariant j. Adv. in Math. 35 (1980), no. 3, 264–273. MR563927 (83a:17014b)

[KL] Kawahigashi, Yasuyuki; Longo, Roberto. Local conformal nets arising from framed vertex operator algebras. (English summary) Adv. Math. 206 (2006), no. 2, 729–751. MR2263720 (2007m:81141)

[LM] Lam, Ching Hung; Miyamoto, Masahiko. Niemeier lattices, Coxeter elements, and McKay's E_8-observation on the Monster simple group. Int. Math. Res. Not. 2006, Art. ID 35967 MR2219232 (2007b:17041)

[M] Miyamoto, Masahiko. 21 involutions acting on the Moonshine module. (English summary) J. Algebra 175 (1995), no. 3, 941–965. MR1341752 (96h:20034)

[Ra] Ray, U. Automorphic Forms and Lie Superalgebras (Algebra and Applications) Springer; 1 edition (December 2006) ISBN 978-1402050091 MR2286867

[Ry] Ryba, A. J. E. Modular Moonshine? Moonshine, the Monster, and related topics (South Hadley, MA, 1994), 307–336, Contemp. Math., 193, Amer. Math. Soc., Providence, RI, 1996. MR1372729 (97c:20022)

[S] Smith, Stephen D. On the head characters of the Monster simple group. Finite groups—coming of age (Montreal, Que., 1982), 303–313, Contemp. Math., 45, Amer. Math. Soc., Providence, RI, 1985. MR822245 (87h:20037)

[T] Thompson, J. G. Some numerology between the Fischer-Griess Monster and the elliptic modular function. Bull. London Math. Soc. 11 (1979), no. 3, 352–353. MR554402 (81j:20030)

[W] Witten, E. Three-Dimensional Gravity Revisited, arXiv:0706.3359

R. E. BORCHERDS
UNIVERSITY OF CALIFORNIA AT BERKELEY

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

arXiv:1702.03516v4 [math.NT] 17 Mar 2019

O'NAN MOONSHINE AND ARITHMETIC

JOHN F. R. DUNCAN, MICHAEL H. MERTENS, AND KEN ONO

Abstract. Answering a question posed by Conway and Norton in their seminal 1979 paper on moonshine, we prove the existence of a graded infinite-dimensional module for the sporadic simple group of O'Nan, for which the McKay–Thompson series are weight $3/2$ modular forms. The coefficients of these series may be expressed in terms of class numbers, traces of singular moduli, and central critical values of quadratic twists of weight 2 modular L-functions. As a consequence, for primes p dividing the order of the O'Nan group we obtain congruences between O'Nan group character values and class numbers, p-parts of Selmer groups, and Tate–Shafarevich groups of certain elliptic curves. This work represents the first example of moonshine involving arithmetic invariants of this type.

1. Introduction and Statement of Results

The sporadic simple groups are the twenty-six exceptions to the classification [7] of finite simple groups: those examples that aren't included in any of the natural infinite families. It is natural to wonder where they appear, outside of the classification itself.

At least for the *monster*, being the largest of the sporadics, this question has an interesting answer. By the last decade of the last century, Ogg's observation [72] on primes dividing the order of the monster, McKay's famous formula

$$196884 = 1 + 196883,$$

and the much broader family of coincidences observed by Thompson [89, 90] and Conway–Norton [29], were proven by Borcherds [10] to reflect the existence of a certain distinguished algebraic structure. This *moonshine module*, constructed by Frenkel–Lepowsky–Meurman [44, 45, 46], admits a vertex operator algebra structure, has the monster as its full symmetry group, and has modular functions for traces. It is a cornerstone of *monstrous moonshine*, and indicates a pathway by which ideas from theoretical physics, and string theory in particular, may ultimately reveal a natural origin for the monster group and its curious connection to modularity.

In addition to the monster itself, nineteen of the sporadic simple groups appear as quotients of subgroups of the monster. As such, we may expect that monstrous moonshine extends to them in some form. This is consequent upon the *generalized moonshine conjecture*, which was formulated by Norton [71] following preliminary observations of Conway–Norton [29] and Queen [78], and which has been recently proven in powerful work by Carnahan [21].

Certain more general analogues of monstrous moonshine have appeared in this century. In 2010, Eguchi–Ooguri–Tachikawa [41] sparked a resurgence of interest in moonshine with their observation that the elliptic genus of a K3 surface—a trace function arising from a non-linear

---

2010 *Mathematics Subject Classification.* 11F22, 11F37.

sigma model with K3 target—is, essentially, the sum of an indefinite theta function and a q-series whose coefficients are dimensions of modules for Mathieu's largest sporadic group, M_{24}. In fact, this q-series is a mock modular form which, together with most of Ramanujan's mock theta functions, belongs to a family of distinguished examples [24] arising from a family of finite groups. This is *umbral moonshine* [26, 27, 28], and the existence of corresponding umbral moonshine modules has been verified by Gannon [49] in the case of M_{24}, and in general by Griffin and two of the authors of this work [37]. However, it must be noted that this theory is not yet on the same footing as monstrous moonshine, as suitable umbral counterparts to the moonshine module vertex operator algebra of Frenkel–Lepowsky–Meurman are not yet known in general.

We refer the reader to [46, 47] for fuller discussions of monstrous moonshine, and to Gannon's book [48] for a broad perspective on the theory. The more recent review [38] includes some umbral developments. We refer to [75, 76] for new work on the string theoretic interpretation of monstrous moonshine, and refer to [6, 25, 39, 40] for vertex algebraic constructions of some of the umbral moonshine modules.

Very recently, yet another form of moonshine has appeared in work of Harvey–Rayhaun [58] which manifests a kind of half-integral weight counterpart to generalized moonshine for Thompson's sporadic group. This is known as *Thompson moonshine*. The existence of a corresponding module has been confirmed by Griffin and one of the authors [54] (but in this case too, a vertex algebraic realization is yet to be found).

All the umbral groups are involved in the monster in some way, so we are left to wonder if there are counterparts to monstrous moonshine for the remaining six *pariah* sporadic groups: the *Janko groups* J_1, J_3, and J_4, the *Lyons group Ly*, the *Rudvalis group Ru*, and the *O'Nan group O'N*. Can moonshine shed light on these groups too? Conway and Norton asked this question (cf. p. 321 of [29]) in their seminal 1979 paper:

"Finally, we ask whether the sporadic simple groups that may not be involved in [the monster]... have moonshine properties."

This question is also Problem #9 in the 1998 paper by Borcherds entitled "Problems in Moonshine" [12].

Rudvalis group analogues of the moonshine module were constructed in [34, 35], but the physical significance of these structures is yet to be illuminated. In this work we present a new form of moonshine which reveals a role for the O'Nan group in arithmetic: as an organizing object for congruences between class numbers, p-parts of Selmer groups and Tate–Shafarevich groups of elliptic curves. (See e.g. [84] or [88] for background on elliptic curve arithmetic.) This is the first occurrence of moonshine of this type. Since J_1 is a subgroup of $O'N$ it suggests that at least two pariah groups play an active part in some of the deepest open questions in arithmetic.

1.1. **Moonshine and Divisors.** Before describing our results in more detail we offer a conceptual number theoretic perspective which ties together some of the recent developments mentioned above. Suppose that G is one of the finite groups appearing in the aforementioned cases of moonshine. Then we have an infinite-dimensional graded G-module, say V^G, which manifests a collection of modular forms, one for each conjugacy class. For monstrous, umbral, and Thompson moonshine we have

$$V^G = \bigoplus_m V_m^G \xrightarrow{\text{moonshine}} (f_{[g]}) \in \begin{cases} \bigoplus\limits_{[g]\in \mathrm{Conj}(G)} M_0^!(\Gamma_{[g]}) & \text{monstrous} \\ \\ \bigoplus\limits_{[g]\in \mathrm{Conj}(G)} H_{\frac{1}{2}}(\Gamma_{[g]}) & \text{umbral, Thompson.} \end{cases}$$

The defining feature of the $f_{[g]}$ is that their m^{th} coefficients equal the graded traces $\mathrm{tr}(g|V_m^G)$.

In monstrous moonshine, the $f_{[g]}$ are Hauptmoduln for genus 0 groups $\Gamma_{[g]}$ (essentially level $o(g)$ congruence subgroups). At the cusp ∞, they have Fourier expansion

$$f_{[g]} = q^{-1} + O(q)$$

(note $q := e^{2\pi i\tau}$ throughout), and are holomorphic at other cusps. In particular, this means that $\mathrm{div}(f_{[g]}) = cz - \infty$ for some $z \in X(\Gamma_{[g]})$ and some integer c. In contrast, the $f_{[g]}$ in umbral and Thompson moonshine are not functions on modular curves, so it does not generally make sense to consider their divisors. Instead, they are weight $1/2$ harmonic Maass forms (with multiplier) for $\Gamma_{[g]}$, which means that the McKay–Thompson series are generally *mock modular forms*, the holomorphic parts of the $f_{[g]}$. Although they are not functions on these modular curves, it turns out that they actually encode even more information about divisors on $X(\Gamma_{[g]})$. For each discriminant D, there is a map Ψ_D for which

$$V^G = \bigoplus_m V_m^G \xrightarrow{\text{moonshine}} (f_{[g]}) \xrightarrow{\Psi_D} (\Psi_D(f_{[g]})) \in \bigoplus_{[g]\in \mathrm{Conj}(G)} \mathcal{K}(\Gamma_{[g]}),$$

where $\mathcal{K}(\Gamma_{[g]})$ is the field of modular functions for $\Gamma_{[g]}$. The $\Psi_D(f_{[g]})$ are *generalized Borcherds products* as defined by Bruinier and one of the authors [20]. They are meromorphic modular functions with a discriminant D Heegner divisor, and their fields of definition are dictated by the Fourier coefficients of the $f_{[g]}$.

As the preceding discussion illustrates, monstrous, umbral, and Thompson moonshine are (surprising) phenomena in which a single infinite-dimensional graded G-module organizes information about divisors on products of modular curves that are indexed by the conjugacy classes of G. Moreover, the levels of these modular curves are (essentially) the orders of elements in these classes. In the case of monstrous moonshine, the divisors are simple: they are of the form $cz - \infty$. In umbral and Thompson moonshine, we obtain Heegner divisors on $X(\Gamma_{[g]})$.

The appearance of Heegner divisors recalls the seminal work of Zagier [99] on traces of singular moduli on $X_0(1)$. Loosely speaking, Zagier proved that the generating function for such traces in D-aspect can be weight $3/2$ weakly holomorphic modular forms. One of his motivations was to offer a classical perspective on special cases of Borcherds' work [11] on infinite product expansions of modular forms with Heegner divisor.

Although Zagier's paper has inspired too many papers to mention, we highlight an important note by Gross [55]. Gross observed that these types of theorems could be recast in terms of *generalized Jacobians* with cuspidal moduli. In particular, the generalized Jacobian of $X_0(1)$ with respect to the cuspidal divisor $2(\infty)$ is isomorphic to the additive group, and so the sum of the conjugates of Heegner points in the generalized Jacobian is equal to the trace of their modular invariants.

Here we adopt this perspective. Although we do not directly apply these results in this work, our view is that the McKay–Thompson series presented here should be viewed in this way, as generating functions for traces of singular moduli and as functionals on Heegner divisors. This interpretation is an extension of the celebrated theorem of Gross–Kohnen–Zagier [56] which asserts that the generating function for Heegner divisors on $X_0(N)$ are weight 3/2 cusp forms with values in the Jacobian of $X_0(N)$. By work of Waldspurger [93, 94] this earlier theorem can be thought of as a result on central critical values of quadratic twists of weight 2 modular L-functions.

1.2. **Main Results.** In view of these developments, it is natural to seek weight 3/2 moonshine. One can loosely think of this as the moonshine obtained by summing weight 1/2 moonshine in D-aspect (e.g. umbral and Thompson moonshine), where the resulting McKay–Thompson series are generating functions for the arithmetic of Heegner divisors. Namely, we seek moonshine of the form

$$V^G = \bigoplus_m V_m^G \quad \xrightarrow{\text{moonshine}} \quad (f_{[g]}) \in \bigoplus_{[g]\in\mathrm{Conj}(G)} H_{\frac{3}{2}}(\Gamma_{[g]}) \otimes \mathrm{Jac}(X(\Gamma_{[g]})),$$

where $\mathrm{Jac}(X(\Gamma_{[g]})$ denotes a suitable generalized Jacobian of $X(\Gamma_{[g]})$. In such moonshine, the $f_{[g]}$ will be generating functions for suitable functionals over Heegner divisors. Their coefficients will be sums of class numbers, traces of singular moduli, and square-roots of central critical values of L-functions of quadratic twists of weight 2 modular forms.

Here we establish the first example of moonshine of this type, and it is pleasing that pariah sporadic groups appear. We prove moonshine for the O'Nan group *O'N*, a group discovered in 1973 as part of the flurry of activity related to the classification of finite simple groups [73] and shown not to be involved in the monster by Griess [52, Lemma 14.5]. This group was first constructed by Sims (cf. [73, p. 421]), and Ryba [81] later gave an alternative construction. It has order $\#O'N = 2^9\cdot 3^4\cdot 5\cdot 7^3\cdot 11\cdot 19\cdot 31$, and it has 30 conjugacy classes. It contains the first Janko group J_1, also not involved in the monster [95], as a subgroup.

Theorem 1.1. *There is an infinite-dimensional virtual graded O'N-module*

$$W := \bigoplus_{0<m\equiv 0,3 \pmod 4} W_m$$

and weight 3/2 modular forms $\{F_{1A}, F_{2A}, \ldots, F_{31A}, F_{31B}\}$*, one for each conjugacy class, with the property that*

$$F_{[g]}(\tau) = -q^{-4} + 2 + \sum_{0<m\equiv 0,3 \pmod 4} \mathrm{tr}(g|W_m)q^m.$$

Moreover, each $F_{[g]}$ *is on the group* $\Gamma_0(4o(g))$*, with a non-trivial character in case* $o(g) = 16$*, and satisfies the Kohnen plus space condition.*

Remark. There is an alternative to Theorem 1.1 in which the $F_{[g]}$ have trivial characters for all g, but are mock modular for $o(g) = 16$. That formulation featured in an earlier version of this work. The present statement is motivated by cohomological considerations and related structures in the representation theory of vertex operator algebras, as we explain in more detail in Section 3. We also characterize the $F_{[g]}$ precisely in Section 3.

Remark. In other prominent examples of moonshine (e.g. monstrous [29] and umbral [26, 27] moonshine) the McKay–Thompson series of a group element g is a modular form (essentially) of level $o(g)$, but in this work the McKay–Thompson series $F_{[g]}$ have level $4o(g)$. This anomaly can be resolved by repackaging the $F_{[g]}$ as Jacobi forms as follows. For $g \in O'N$ set

$$\varphi_{[g]}(\tau, z) := F_{[g],0}(\tfrac{\tau}{4})\theta_{1,0}(\tau, z) + F_{[g],1}(\tfrac{\tau}{4})\theta_{1,1}(\tau, z) \tag{1.1}$$

where $F_{[g],r}(\tau) := \sum_{m\equiv r \bmod 2} \operatorname{tr}(g|W_m)q^m$ and $\theta_{1,r}(\tau, z) := \sum_{n\equiv r \bmod 2} e^{2\pi i n z} q^{\frac{n^2}{4}}$. Then $\varphi_{[g]}$ is a weakly holomorphic Jacobi form of weight 2 and index 1 on $\Gamma_0(o(g))$, with a non-trivial character in case $o(g) = 16$. For the sake of simplicity we have chosen to formulate our results in terms of the scalar-valued modular forms $F_{[g]}$ in this work. However, we note that one advantage of the Jacobi form formulation is that it illuminates an analogue of the Hauptmodul property of monstrous moonshine. Namely, each $\varphi_{[g]}$ has the property that it is uniquely determined, up to a cusp form, by the condition that it has growth of a certain form (independent of g) near the infinite cusp of $\Gamma_0(o(g))$, and vanishes at all other cusps. This follows from the proof of Theorem 3.1. It may be compared to monstrous moonshine, in which the McKay–Thompson series are uniquely determined up to constant functions by an analogous condition, and to umbral and Thompson moonshine, in which the McKay–Thompson series are uniquely determined by such a condition up to theta series (although in the umbral case almost all the relevant spaces of theta series vanish; cf. [28]). It is the appearance of cusp forms that allows us to connect the O'Nan group to elliptic curve arithmetic.

Remark. The module W is *virtual* in the sense that some irreducible representations of $O'N$ occur with negative multiplicity in W_m for some m. The proof of Theorem 1.1 will show that only non-negative multiplicities appear for $m \notin \{7, 8, 12, 16\}$. So in fact we can replace W with a non-virtual module for a small cost, by adding suitable multiples of weight 3/2 unary theta functions (i.e. sums of the form $\sum_{n\in\mathbb{Z}} n\epsilon(n)q^{\lambda n^2}$ where ϵ is an odd periodic function and λ is a positive rational) to the McKay–Thompson series $F_{[g]}$. This changes the module structure of W_m when $m = vd^2$ for $v \in \{7, 8, 12, 16\}$, for certain integers d, but it does not effect the validity of our other three main results, Theorems 1.2, 1.3, and 1.4, for $-D < -16$. The price for such an adjustment to W is the property that the McKay–Thompson series attached to $[g]$ have level $4o(g)$. It is this property which motivates us to focus on the particular module W that appears in Theorem 1.1.

The $F_{[g]}$ will turn out to be expressible in terms of traces of singular moduli for Hauptmoduln (cf. Section 5), class numbers, and central critical L-values of quadratic twists of weight 2 modular forms (cf. Section 4.2.2). The Hauptmoduln which arise are for the genus 0 modular curves

$$\{X_0(N) : N = 1, \dots, 8, 10, 12, 16\} \cup \{X_0^+(N) : N = 11, 14, 15, 16, 19, 20, 28, 31, 32\}, \tag{1.2}$$

where $X_0^+(N)$ is the modular curve corresponding to the extension of $\Gamma_0(N)$ by all the level N Atkin–Lehner involutions.

Remark. Purely for the sake of curiosity we mention that it follows from the description of the dimensions of the graded components W_m in terms of traces of singular moduli (cf. Appendix D) that

$$\dim W_{163} = \frac{1}{2}(\alpha^2 + \alpha - 393768),$$

where

$$\alpha = \left\lceil e^{\pi\sqrt{163}} \right\rceil = \lceil 262537412640768743.99999999999250072... \rceil$$

denotes the Ramanujan constant. (This number was actually already discovered and studied by Hermite in 1859 [59].)

Remark. From Tables B.1 to B.3 we see that W_3 is an irreducible *O'N*-module of dimension 26752, and W_4 has three irreducible constituents, with dimensions 1, 58311 and 85064. On the other hand the specialization $\varphi_{1A}(\tau, 0)$ of Equation (1.1) is the derivative of the J function, up to a scalar factor. This leads to the identity

$$196884 = 5 \cdot 1 + 2 \cdot 26752 + 58311 + 85064,$$

where the summands on the right are dimensions of irreducible representations of *O'N*. Inspired by the moonshine module vertex operator algebra [46] of Frenkel–Lepowsky–Meurman we may ask: is there a holomorphic vertex operator algebra with an action by *O'N* that explains this coincidence? (See the second remark in Section 3 for some further related comments.)

Armed with Theorem 1.1 and the explicit identities expressing the $F_{[g]}$ in terms of singular moduli, class numbers and critical L-values, it is natural to ask whether the infinite-dimensional *O'N*-module W reveals arithmetic information about the modular curves they organize, which include the positive genus curves

$$\{X_0(11),\ X_0(14),\ X_0(15),\ X_0(19),\ X_0(20),\ X_0(28),\ X_0(31)\}$$

related to the $X_0^+(N)$ in (1.2). For example, are there interesting congruences modulo primes $p|\#O'N$ which relate the graded components W_m to classical objects in number theory and arithmetic geometry? This is indeed the case, and we now describe surprising congruences which relate graded dimensions and traces of W to class numbers and Selmer groups and Tate–Shafarevich groups of elliptic curves.

Remark. Suppose that p is prime and g_n (resp. g_{np}) are elements of *O'N* with order n (resp. np). Then by Theorem 1.1, we have that $\operatorname{tr}(g_n|W_m) \equiv \operatorname{tr}(g_{np}|W_m) \pmod{p}$ for all m. In particular, if $o(g) = p$, then for all m we have

$$\dim W_m \equiv \operatorname{tr}(g|W_m) \pmod{p}.$$

The following theorem concerns congruences modulo small primes p and ideal class groups of imaginary quadratic fields. Here and in the following, we denote by $H(D)$ the Hurwitz class number of positive definite binary quadratic forms of discriminant $-D < 0$ (cf. Section 5).

Theorem 1.2. *Suppose that $-D < 0$ is a fundamental discriminant. Then the following are true:*

(1) *If* $-D < -8$ *is even and* $g_2 \in O'N$ *has order 2, then*

$$\dim W_D \equiv \operatorname{tr}(g_2|W_D) \equiv -24H(D) \equiv 0 \pmod{2^4}.$$

(2) *If* $p \in \{3, 5, 7\}$, $\left(\frac{-D}{p}\right) = -1$ *and* $g_p \in O'N$ *has order* p, *then*

$$\dim W_D \equiv \operatorname{tr}(g_p|W_D) \equiv \begin{cases} -24H(D) \pmod{3^2} & \text{if } p = 3, \\ -24H(D) \pmod{p} & \text{if } p = 5, 7. \end{cases}$$

Remark. Systematic congruences which assert for $\left(\frac{-D}{p}\right) = -1$ that

$$\dim W_D \equiv -24H(D) \pmod{p}$$

do not seem to hold for $p \geq 17$. However, this congruence holds for $p = 13$, a bonus because $13 \nmid \#O'N$.

Remark. As the proof of Theorem 1.2 will reveal, it holds true that if $-D < -8$ is an even fundamental discriminant, then $H(D)$ is even, and $\dim W_D \equiv 0 \pmod{2^4}$.

In view of Theorem 1.2, it is natural to consider the primes $p = 11, 19$ and 31 which also divide $\#O'N$. For these primes, a refinement of the congruences above is necessary. In particular, for the primes 11 and 19 we obtain congruences which relate $\dim W_D$ to Selmer groups and Tate–Shafarevich groups of elliptic curves (cf. [84, Chapter X]).

Let $E/\mathbb{Q}$ be an elliptic curve given by

$$E: \quad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

where $a_1, a_2, a_3, a_4, a_6 \in \mathbb{Z}$. For a fundamental discriminant D, let $E(D)$ denote its D-quadratic twist, and let $\operatorname{rk}(E(D))$ denote the Mordell–Weil rank of $E(D)$ over $\mathbb{Q}$. The $O'N$-module W encodes deep information about the Selmer and Tate–Shafarevich groups of the quadratic twists of elliptic curves with conductor 11, 14, 15, and 19. To make this precise, suppose that ℓ is an odd prime. Then for each curve $E(D)$ we have the short exact sequence

$$1 \to E(D)/\ell E(D) \to \operatorname{Sel}(E(D))[\ell] \to Ш(E(D))[\ell] \to 1,$$

where $\operatorname{Sel}(E(D))[\ell]$ is the ℓ-Selmer group of $E(D)$, and $Ш(E(D))[\ell]$ denotes the elements of the Tate–Shafarevich group $Ш(E(D))$ with order dividing ℓ.

For $p = 11$ and 19, we let $E_p/\mathbb{Q}$ be the $\Gamma_0(p)$-optimal elliptic curves given by the Weierstrass models

$$E_{11}: \quad y^2 + y = x^3 - x^2 - 10x - 20,$$
$$E_{19}: \quad y^2 + y = x^3 + x^2 - 9x - 15$$

(cf. [68, Elliptic Curve 11.a2, Elliptic Curve 19.a2]). We obtain the following congruence relating the graded dimension $\dim W_D$ to class numbers, and Selmer groups and Tate–Shafarevich groups of such twists.

Theorem 1.3. *Assume the Birch and Swinnerton-Dyer Conjecture. If* $p = 11$ *or* 19 *and* $-D < 0$ *is a fundamental discriminant for which* $\left(\frac{-D}{p}\right) = -1$, *and* $g_p \in O'N$ *has order* p, *then the following are true.*

(1) We have that $\mathrm{Sel}(E_p(-D))[p] \neq \{0\}$ *if and only if*

$$\dim W_D \equiv \mathrm{tr}(g_p|W_D) \equiv -24H(D) \pmod{p}.$$

(2) Suppose that $L(E_p(-D),1) \neq 0$. *Then we have that* $\mathrm{rk}(E(-D)) = 0$. *Moreover, we have* $p|\#\text{Ш}(E_p(-D))$ *if and only if*

$$\dim W_D \equiv \mathrm{tr}(g_p|W_D) \equiv -24H(D) \pmod{p}.$$

Remark. The claim about ranks in Theorem 1.3 (2) is unconditional thanks to the work of Kolyvagin [67].

Remark. By Goldfeld's famous conjecture on ranks of quadratic twists of elliptic curves [51], it turns out that the hypothesis in Theorem 1.3 (2) is expected to hold for 100% of the $-D$ for which $\left(\frac{-D}{p}\right) = -1$. Therefore, for almost all such $-D$, we should have a test for determining the presence of order p elements in these Tate–Shafarevich groups.

Remark. There is a more complicated congruence for the prime $p = 31$. For fundamental discriminants $-D < 0$ satisfying $\left(\frac{-D}{31}\right) = -1$, we have that $\dim W_D \equiv \mathrm{tr}(g_{31}|W_D) \pmod{31}$ are related to the central critical values of the $-D$ twists of the L-function for the genus 2 curve

$$C: \quad y^2 + (x^3+x+1)y = x^5 + x^4 + x^3 - x - 1$$

(cf. [68, Genus 2 Curve 961.a.961.3]). Its L-function arises from the two newforms in $S_2(\Gamma_0(31))$ which are Galois conjugates. Namely, if $\phi := \frac{1+\sqrt{5}}{2}$ then the two newforms are f^σ and

$$f(\tau) := \sum_{n=1}^{\infty} a(n)q^n = q + \phi q^2 - 2\phi q^3 + (\phi-1)q^4 + q^5 - (2\phi+2)q^6 + O(q^7),$$

where $\sigma(\sqrt{5}) = -\sqrt{5}$. If $p \nmid 31$ is prime, then the local L-factor $L_p(T)$ at p is

$$L_p(T) := (1 - a(p)T + pT^2)(1 - \sigma(a(p))T + pT^2).$$

Remark. Apart from the claims about $\mathrm{tr}(g_{17}|W_D)$ (there are no elements of order 17 in $O'N$), Theorem 1.3 holds for $p = 17$ as well. Namely, the congruences hold for E_{17}, the optimal $\Gamma_0(17)$ elliptic curve over $\mathbb{Q}$ (cf. [68, Elliptic Curve 17.a3]) given by

$$E_{17}: \quad y^2 + xy + y = x^3 - x^2 - x - 14.$$

The two theorems on congruences above only pertain to the dimensions of the graded components of the $O'N$-module W. We now turn to congruences for graded traces for elements of order 2 and 3. To this end, we let E_{14} and E_{15} be the corresponding optimal elliptic curves over $\mathbb{Q}$ (cf. see [68, Elliptic Curve 14.a6, Elliptic Curve 15.a5]) given by

$$E_{14}: \quad y^2 + xy + y = x^3 + 4x - 6,$$
$$E_{15}: \quad y^2 + xy + y = x^3 + x^2 - 10x - 10.$$

Using work of Skinner and Skinner–Urban [85, 86] related to the Iwasawa main conjectures for GL_2, we obtain the following unconditional result.

Theorem 1.4. *Assume the notation above, and suppose that* $N \in \{14, 15\}$*. If* p *is the unique prime* ≥ 5 *dividing* N*, then let* $\delta_p := \frac{p-1}{2}$ *and let* $p' := N/p$*. If* $-D < 0$ *is a fundamental discriminant for which* $\left(\frac{-D}{p}\right) = -1$ *and* $\left(\frac{-D}{p'}\right) = 1$*, then the following are true.*

(1) We have that $\mathrm{Sel}(E_N(-D))[p] \neq \{0\}$ *if and only if*

$$\mathrm{tr}(g_{p'}|W_D) \equiv \mathrm{tr}(g_N|W_D) \equiv \delta_p \cdot (H(D) - \delta_p H^{(p')}(D)) \pmod{p}.$$

(2) Suppose that $L(E_N(-D), 1) \neq 0$*. Then we have that* $\mathrm{rk}(E(-D)) = 0$*. Moreover, we have* $p|\#\text{Ш}(E_N(-D))$ *if and only if*

$$\mathrm{tr}(g_{p'}|W_D) \equiv \mathrm{tr}(g_N|W_D) \equiv \delta_p \cdot (H(D) - \delta_p H^{(p')}(D)) \pmod{p}.$$

Remark. We note that Theorem 1.4 does not apply for $p = 2$ (resp. $p = 3$) when $N = 14$ (resp. $N = 15$). In the case of $p = 2$ the work of Skinner–Urban does not apply. For $p = 3$ the connection between graded traces and central values of Hasse-Weil L-functions does not hold. Namely, a critical hypothesis due to Kohnen in terms of eigenvalues of Atkin–Lehner involutions fails (cf. Proposition 4.4).

Remark. In view of the new results presented here, it is natural to wonder where one should look for further moonshine. It seems likely that other sporadic groups will fall within the scope of weight $3/2$ moonshine. In another direction, one can ask about other half-integral weights. Also, it is natural to wonder if there are extensions of moonshine to Shimura curves and varieties. Are there infinite-dimensional G-modules which organize the arithmetic of their divisors?

1.3. **Methods.**

To prove Theorem 1.1, we employ the theory of Rademacher sums, harmonic Maass forms, and standard facts about the representation theory of finite groups. Namely, we make use of the character table of *O'N* (cf. Table A.1), and the Schur orthogonality relations for group characters. In Section 2, we first recall essential facts about harmonic Maass forms and Rademacher sums. In Section 3, we prove a theorem which, using harmonic Maass forms, explicitly constructs weakly holomorphic weight $3/2$ modular forms, one for each conjugacy class of *O'N*. Furthermore, we establish that these modular forms have integer Fourier coefficients. To complete the proof, we apply the Schur orthogonality relations to these functions to construct weight $3/2$ modular forms whose coefficients encode the multiplicities of the irreducibles of the graded components of the alleged module W. The proof is complete once it is established that these multiplicities are integral. Since the obstruction to integrality is bounded by group theoretical considerations, the proof of integrality follows by confirming sufficiently many congruence relations among these forms. These calculations confirm that W is a virtual module. However, as mentioned earlier, it turns out that the multiplicities of each irreducible are non-negative in W_m once $m > 16$. This claim follows from an analytic argument which involves bounding sums of Kloosterman sums. These statements are proved in Section 4. In Section 5 we recall properties of singular moduli, and we interpret the modular forms number theoretically in terms of singular moduli and class numbers and cusp forms. We prove Theorems 1.2, 1.3 and 1.4 in Section 6. These proofs require the explicit formulas for the $F_{[g]}$, the results in Section 5, and the work of Skinner–Urban on the Birch and Swinnerton-Dyer Conjecture. We conclude the paper in Section 7 with numerical examples of some of these results.

Acknowledgements

The authors thank Kathrin Bringmann, Michael Griffin, Dick Gross, Maryam Khaqan, Winfried Kohnen, Martin Raum, Jeremy Rouse, Jean-Pierre Serre and anonymous referees for helpful comments and corrections. The authors thank Theo Johnson-Freyd for communication regarding his joint work with David Treumann on the cohomology of the O'Nan group. The authors thank Drew Sutherland for computing the elliptic curve invariants in Tables 7.5 and 7.6.

2. Rademacher Sums and Harmonic Maass Forms

Harmonic Maass forms are now a central topic in number theory. Their study originates from the work of Bruinier–Funke [18] on geometric theta lifts and Zwegers' seminal work [101] on Ramanujan's mock theta functions. These realizations played a central role in the work of Bringmann and one of the authors on the Andrews–Dragonette Conjecture and Dyson's partition ranks [15, 17]. For an overview on the subject of harmonic Maass forms and its applications in number theory and various other fields of mathematics, including mathematical physics, we refer the reader to [14, 31, 74, 100].

Here, we briefly recall the essential facts about harmonic Maass forms that are required in this paper. Namely, we recall Rademacher sums, and we describe their projection to Kohnen's plus space.

2.1. **Rademacher Sums.** Here and throughout, we let $\tau = u + iv$, $u, v \in \mathbb{R}$, denote a variable in the upper half-plane $\mathfrak{H}$ and we use the shorthand $e(\alpha) := e^{2\pi i \alpha}$.

Definition 2.1. *We call a smooth function $f : \mathfrak{H} \to \mathbb{C}$ a harmonic Maass form of* weight $k \in \frac{1}{2}\mathbb{Z}$ *and* level N *if the following conditions are satisfied:*

(1) *We have $f|_k\gamma(\tau) = f(\tau)$ for all $\gamma \in \Gamma_0(N)$ and $\tau \in \mathfrak{H}$, where we define*

$$f|_k\gamma(\tau) := \begin{cases} (c\tau+d)^{-k} f\left(\frac{a\tau+b}{c\tau+d}\right) & \text{if } k \in \mathbb{Z} \\ \left(\left(\frac{c}{d}\right)\varepsilon_d\right)^{2k} \left(\sqrt{c\tau+d}\right)^{-2k} f\left(\frac{a\tau+b}{c\tau+d}\right) & \text{if } k \in \frac{1}{2}+\mathbb{Z}. \end{cases}$$

with

$$\varepsilon_d := \begin{cases} 1 & \text{if } d \equiv 1 \pmod 4, \\ i & \text{if } d \equiv 3 \pmod 4. \end{cases}$$

and where we assume $4|N$ if $k \notin \mathbb{Z}$.

(2) *The function f is annihilated by the* weight k hyperbolic Laplacian,

$$\Delta_k f := \left[-v^2\left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}\right) + ikv\left(\frac{\partial}{\partial u} + i\frac{\partial}{\partial v}\right)\right] f \equiv 0.$$

(3) *There is a polynomial $P(q^{-1})$ such that $f(\tau) - P(e^{-2\pi i\tau}) = O(v^c)$ for some $c \in \mathbb{R}$ as $v \to \infty$. Analogous conditions are required at all cusps of $\Gamma_0(N)$.*

We denote the space of harmonic Maass forms of weight k and level N by $H_k(\Gamma_0(N))$.

Remark. We note that condition (3) in the definition above differs from other definitions which occur commonly in the literature. For example, *harmonic Maass forms with principal parts* are those forms for which the $O(v^c)$ bound is replaced by $O(e^{-cv})$ for $c > 0$. Namely, the harmonic Maass forms we consider here are permitted to have 0^{th} Fourier coefficients which are essentially powers of v.

For the basic properties of these functions, we again refer to the literature mentioned above. We mention however the following lemmas.

Lemma 2.2. *Let $f \in H_k(\Gamma_0(N))$ be a harmonic Maass form of weight $k \neq 1$. Then there is a canonical splitting*

$$f(\tau) = f^+(\tau) + f^-(\tau), \tag{2.1}$$

where for some $m_0 \in \mathbb{Z}$ we have the Fourier expansions

$$f^+(\tau) := \sum_{n=m_0}^{\infty} c_f^+(n) q^n,$$

and

$$f^-(\tau) := \overline{c_f^-(0)} \frac{(4\pi v)^{1-k}}{k-1} + \sum_{n=1}^{\infty} \overline{c_f^-(n)} n^{k-1} \Gamma(1-k; 4\pi n v) q^{-n},$$

where

$$\Gamma(\alpha; x) := \int_x^{\infty} t^{\alpha} e^{-t} \frac{dt}{t}$$

denotes the usual incomplete gamma function.

The q-series f^+ in (2.1) is called the *holomorphic part* of the harmonic Maass form f.

An important differential operator in the theory of harmonic Maass forms is the ξ-operator, a variation of the Maass lowering operator.

Proposition 2.3. *The operator*

$$\xi_k : H_k(\Gamma_0(N)) \to M_{2-k}(\Gamma_0(N)),\ f \mapsto \xi_k(f) := 2iy^k \overline{\frac{\partial f}{\partial \overline{\tau}}}$$

is a well-defined and surjective anti-linear map with kernel $M_k^!(\Gamma_0(N))$.

Mock modular forms are the holomorphic parts of harmonic Maass forms. Any mock modular form has an associated modular form, called its *shadow*, which is the image of its corresponding harmonic Maass form under the ξ-operator. A mock modular form with vanishing shadow is a (weakly holomorphic) modular form.

The next lemma seems to have been missed by the literature.

Lemma 2.4. *A harmonic Maass form whose holomorphic part vanishes at all cusps is a (holomorphic) cusp form.*

Proof. This is a direct consequence of the properties of the Bruinier–Funke pairing (cf. Proposition 3.5 in [18]). □

A convenient way to construct mock modular forms, which are holomorphic parts of harmonic Maass forms, is through *Rademacher sums*. These were introduced by Rademacher in his work on coefficients of the J-function [79], and further developed in the context of moonshine mainly by Cheng, Frenkel and one of the authors [22, 23, 36].

Rademacher sums can be thought of as low weight analogues of Poincaré series. For a fixed level N and some $K > 0$, we define the set

$$\Gamma_{K,K^2}(N) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N) \; : \; |c| < K \text{ and } |d| < K^2 \right\}.$$

Given an integer μ we can use this to formally define the Rademacher sum

$$R^{[\mu]}_{k,N}(\tau) := \lim_{K\to\infty} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_{K,K^2}(N)} q^\mu|_k\gamma$$

where as usual $\Gamma_\infty := \{\pm \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right) \; : \; n \in \mathbb{Z}\}$ denotes the stabilizer of ∞ in $\Gamma_0(N)$. If convergent, these sums define mock modular forms of the indicated weight, level and character. Convergence for these series however is in general a delicate matter when the weight k is between 0 and 2. We will be interested in these series when the weight is $k = \frac{3}{2}$ in which case it has been established in [22, Section 5] that they do converge (possibly using a certain regularization explained in loc. cit.) and define holomorphic functions on $\mathfrak{H}$.

By construction, Rademacher sums are 1-periodic and therefore have a Fourier expansion. It is given in terms of infinite sums of *Kloosterman sums*

$$K_k(m,n,c) := \sum_{d \pmod c}^{*} \left(\frac{c}{d}\right) \varepsilon_d^{2k} e\left(\frac{m\overline{d}+nd}{c}\right) \tag{2.2}$$

weighted by Bessel functions. Here we have that $k \in \frac{1}{2} + \mathbb{Z}$, c is divisible by 4, the * at the sum indicates that it runs over primitive residue classes modulo c, and $\overline{d}$ denotes the multiplicative inverse of d modulo c. Computing the Fourier expansion of a Rademacher sum is a standard computation, see for instance [23, Section 3.1] and [74, Section 8.3].

Theorem 2.5. *Assuming locally uniform convergence, for $\mu \leq 0$ and $k \in \frac{1}{2}+\mathbb{N}$ and $4|N$, the Rademacher sum $R^{[\mu]}_{k,N}$ defines a mock modular form of weight k for $\Gamma_0(N)$ whose shadow is given by a constant multiple of the Rademacher sum $R^{[-\mu]}_{2-k,N}$. Its Fourier expansion is given by*

$$R^{[\mu]}_{k,N}(\tau) = q^\mu + \sum_{n=1}^{\infty} c^{[\mu]}_{k,N}(n) q^n,$$

where

$$c^{[\mu]}_{k,N}(n) = -2\pi i^k \left|\frac{n}{\mu}\right|^{\frac{k-1}{2}} \sum_{\substack{c>0 \\ c\equiv 0 \pmod N}} \frac{K_k(\mu,n,c)}{c} \cdot I_{k-1}\left(\frac{4\pi\sqrt{|\mu n|}}{c}\right) \tag{2.3}$$

for $\mu < 0$ and

$$c^{[0]}_{k,N}(n) = (-2\pi i)^k \frac{n^{k-1}}{\Gamma(k)} \sum_{\substack{c>0 \\ c\equiv 0 \pmod N}} \frac{K_k(0,n,c)}{c^k}. \tag{2.4}$$

The completion $\widehat{R^{[\mu]}_{k,N}}$ of $R^{[\mu]}_{k,N}$ to a harmonic Maass form has a pole of order μ at the cusp ∞ and vanishes at all other cusps.

Remark. One can also consider Rademacher sums of weights $\leq 1/2$, which are the main subject of [22] and play a crucial rule in both umbral and Thompson moonshine. The formulas look very similar in those cases, but since they are not needed, we omit them here.

2.2. **Kohnen's Plus Space.** In [66], Kohnen introduced the notion of the so-called plus space, a natural subspace of weight $k + \frac{1}{2}$ cusp forms for $\Gamma_0(4N)$ which is isomorphic via the Shimura correspondence to the space of weight $2k$ cusp forms of level N as a Hecke module, provided that N is odd and square-free. This space is easily characterized via Fourier expansions. Namely, it consists of all forms in $S_{k+\frac{1}{2}}(\Gamma_0(4N))$ (or, by extension, $M^!_{k+\frac{1}{2}}(\Gamma_0(4N))$ and also $H_{k+\frac{1}{2}}(\Gamma_0(4N))$) whose Fourier coefficients are supported on exponents n with $n \equiv 0, (-1)^k \pmod 4$. There is a natural projection operator

$$| \operatorname{pr} : S_{k+\frac{1}{2}}(\Gamma_0(4N)) \to S^+_{k+\frac{1}{2}}(\Gamma_0(4N))$$

for N odd given in terms of slash operators (see loc. cit.), which extends to spaces of weakly holomorphic modular forms and harmonic Maass forms. The action of this projection operator on principal parts of harmonic Maass forms is described in the following lemma (cf. Lemma 2.9 in [54]).

Lemma 2.6. *Let N be odd and let $f \in H_{k+\frac{1}{2}}(\Gamma_0(4N))$ for some $k \in \mathbb{N}_0$. Suppose that*

$$f^+(\tau) = q^{-m} + \sum_{n=0}^{\infty} a_n q^n$$

for some $m > 0$ with $-m \equiv 0, (-1)^k \pmod 4$, and suppose also that f has a non-vanishing principal part only at the cusp ∞ and is bounded at the other cusps of $\Gamma_0(4N)$. Then the projection $f|\operatorname{pr}$ of f to the plus space has a pole of order m at ∞ and has a pole of order $\frac{m}{4}$ either at the cusp $\frac{1}{N}$ if $m \equiv 0 \pmod 4$, or at the cusp $\frac{1}{2N}$ if $-m \equiv (-1)^k \pmod 4$, and is bounded at all other cusps.

For the purpose of this paper, we are particularly interested in the Fourier expansions of weight 3/2 Rademacher sums projected to the plus space (see the following section). Convergence for these follows along the same lines as in [22, Section 5]. The following proposition gives their Fourier expansion explicitly.

Proposition 2.7. *Consider the Rademacher sum $R^{[\mu]}_{\frac{3}{2},4N}$ for $\mu \leq 0$ such that $\mu \equiv 0, 3 \pmod 4$ and N odd. Then we have that*

$$R^{[\mu],+}_{\frac{3}{2},4N}(\tau) := \left(R^{[\mu]}_{\frac{3}{2},4N}|\operatorname{pr}\right)(\tau) = q^{\mu} + \sum_{\substack{n>0 \\ n \equiv 0,3 \pmod 4}} c^{[\mu],+}_{\frac{3}{2},4N}(n) q^n,$$

where we have

$$c^{[\mu],+}_{\frac{3}{2},4N}(n) = \kappa(\mu, n) \sum_{c=1}^{\infty} \left(1 + \delta_{odd}(Nc)\right) K_{\frac{3}{2}}(\mu, n, 4Nc) \cdot \mathcal{I}(\mu, n, 4Nc), \tag{2.5}$$

with

$$\kappa(\mu,n) := \begin{cases} 2\pi e\left(-\frac{3}{8}\right) & \textit{if } \mu=0,\\ 2\pi e\left(-\frac{3}{8}\right)(n/|\mu|)^{\frac{1}{4}} & \textit{otherwise,}\end{cases} \tag{2.6}$$

$$\delta_{odd}(n) := \begin{cases} 1 & \textit{if } n \textit{ is odd,}\\ 0 & \textit{otherwise,}\end{cases} \tag{2.7}$$

and

$$\mathcal{I}(\mu,n,c) := \begin{cases} \dfrac{(2\pi n)^{\frac{1}{2}}}{c^{\frac{3}{2}}\Gamma(3/2)} & \textit{if } \mu=0,\\ \\ \dfrac{I_{\frac{1}{2}}\left(\frac{4\pi\sqrt{|\mu n|}}{c}\right)}{c} & \textit{otherwise.}\end{cases} \tag{2.8}$$

The following proposition shows that the vanishing of Kloosterman sums automatically forces certain even level Rademacher sums to be in the plus space.

Proposition 2.8. *The Rademacher sum* $R^{[\mu]}_{\frac{3}{2},4N}$ *is automatically in the plus space if* N *is even and* $\mu \equiv 0,3 \pmod 4$. *Moreover, if* $N,\mu\equiv 0 \pmod 4$, *then the Fourier coefficients of* $R^{[\mu]}_{\frac{3}{2},4N}$ *are supported on exponents divisible by* 4.

Proof. We begin by noting that if c is divisible by 8, then the Kloosterman sum $K(m,n,c)$ in (2.2) vanishes unless $m-n\equiv 0,3 \pmod 4$. If c is divisible by 16, the same sum vanishes unless $m\equiv n \pmod 4$. Therefore, the claim follows from Theorem 2.5. ☐

Remark. This is an easy restatement (and slight correction) of [54, Lemma 2.10].

Remark. Proposition 2.8 actually follows from the splitting properties of the Weil representation, which is a stronger statement than the vanishing of Kloosterman sums we employed in the proof. However, since we don't use the language of vector-valued modular forms in this paper we use the above more elementary argument.

Remark. We note that the formulas in Proposition 2.7 also hold for even N if one defines the projection operator pr for even levels as a suitable sieving operator, which one easily sees by a comparison to Theorem 2.5.

3. THE RELEVANT MODULAR FORMS

Here we use the results from the previous section to realize the McKay–Thompson series for the O'N-module W whose existence shall be proved later. The main result here is the following theorem. To state it we define a character $\rho_{[g]}:\Gamma_0(4o(g))\to\mathbb{C}^*$ for each conjugacy class $[g]$ of O'N by setting $\rho_{[g]}\left(\begin{smallmatrix} * & * \\ c & * \end{smallmatrix}\right) := (-1)^{\frac{c}{128}}$ when $o(g)=16$, and letting $\rho_{[g]}$ be trivial otherwise.

Theorem 3.1. *Assuming the notation above, the following are true.*

(1) *For every conjugacy class* $[g]$ *of O'N there is a unique weakly holomorphic modular form*

$$F_{[g]}(\tau) = -q^{-4} + 2 + \sum_{n=1} a_{[g]}(n)q^n \tag{3.1}$$

of weight $3/2$ *for the group* $\Gamma_0(4o(g))$*, with character* $\rho_{[g]}$*, satisfying the following conditions:*

  (a) $F_{[g]}(\tau)$ *lies in the Kohnen plus space, i.e.,* $a_{[g]}(n) = 0$ *if* $n \equiv 1, 2 \pmod 4$.

  (b) $F_{[g]}(\tau)$ *has a pole of order* 4 *at the cusp* ∞*, a pole of order* $\frac{1}{4}$ *at the cusp* $\frac{1}{o(g)}$ *if* $o(g)$ *is odd (as forced by the projection to the plus space, see Lemma 2.6), and vanishes at all other cusps.*

  (c) *We have* $a_{[g]}(3) = \chi_7(g)$*, and* $a_{[g]}(4) = \chi_1(g) + \chi_{12}(g) + \chi_{18}(g)$*, and* $a_{[g]}(7)$ *as given in Tables B.1 to B.3, where* χ_j*, for* $j = 1, ..., 30$*, denotes the* j^{th} *irreducible character of O'N as given in Table A.1.*

(2) *The function* $F_{[g]}(\tau)$ *above has integer Fourier coefficients.*

Remark. One can also give a more intrinsic description of the conditions in part (c) above. The proof of the theorem will show that $F_{[g]}$ is already determined by conditions (a) and (b) in part (1), for the 19 conjugacy classes $[g]$ such that $o(g) \notin \{11, 14, 15, 16, 19, 28, 31\}$. For the remaining conjugacy classes we remark that whenever a prime p divides $o(g)$, we need the congruence

$$a_{[g]}(n) \equiv a_{[g']}(n) \pmod p \tag{3.2}$$

where $o(g') = o(g)/p$ in order for these to be generalized characters for *O'N*. Whenever one can choose the coefficient $a_{[g]}(n)$ for the function $F_{[g]}$—which turns out to be the case for $n = 3$ and $o(g) \in \{11, 15, 16\}$, for $n = 4$ for $o(g) \in \{14, 19\}$ and for $n \in \{4, 7\}$ for $o(g) \in \{28, 31\}$—we pick the least integer in absolute value satisfying (3.2) for all primes $p|o(g)$.

Remark. The mod 2 cohomology of the O'Nan group was computed by Adem–Milgram [2]. Using this, Johnson-Freyd–Treumann determined [65] that $H^4(O'N, \mathbb{Z})$ is cyclic of order 8. Furthermore, they have explained to us [64] that there is an element whose image under the restriction map $H^4(O'N, \mathbb{Z}) \to H^4(\langle g\rangle, \mathbb{Z})$ is zero unless $o(g) = 16$, in which case it is the element of order 2 in $H^4(\mathbb{Z}/16\mathbb{Z}, \mathbb{Z}) \simeq \mathbb{Z}/16\mathbb{Z}$. The significance of this is that if V is a holomorphic vertex operator algebra with an action by a finite group G then it is conjectured that the (G-twisted) representation theory of V, including the modularity of its associated trace functions, is controlled by an element of $H^4(G, \mathbb{Z}) \simeq H^3(G, U(1))$. In particular, an element of $H^4(G, \mathbb{Z})$ associates a multiplier system on $\Gamma_0(o(g))$ to each $g \in G$. The above statements about $H^4(O'N, \mathbb{Z})$ imply that there is an element that associates the trivial multiplier to $\Gamma_0(o(g))$ for all $g \in O'N$ except those with $o(g) = 16$, and in the latter case the multiplier arising is the order two character on $\Gamma_0(16)$ with kernel $\Gamma_0(32)$. That is, the characters determined by this element of $H^4(O'N, \mathbb{Z})$ are exactly those that are satisfied by the Jacobi forms $\varphi_{[g]}$ of Equation (1.1), and this is compatible with the existence of a holomorphic vertex operator algebra that realizes these functions, and hence also the $F_{[g]}$, and the *O'N*-module of Theorem 1.1. We refer to §2 of [49], and references therein, for more on the relationship between $H^3(G, U(1))$, modular forms and vertex operator algebra. We refer to §3.2 of [42] for a recent account of the aforementioned conjecture on vertex operator algebras.

Remark. The O'Nan group has a non-split extension $4^3 \cdot \mathrm{GL}_3(2)$ as a subgroup, whilst the sporadic simple Higman–Sims group contains a splitting extension $4^3 : \mathrm{GL}_3(2)$. In both cases the mentioned subgroups contain Sylow 2-subgroups, so can be used to detect the existence, or not, of 2-power elements in the cohomology of the corresponding simple groups. There may be something to be gained from a comparison of these groups, especially in light of the fact that the Higman–Sims group inherits a natural counterpart to monstrous moonshine by virtue of generalized moonshine, and the fact that it appears in the centralizers of suitable elements of order 5 in the monster. We thank an anonymous referee for offering this observation. We also thank the referee for pointing out that the extensions of $\mathrm{GL}_3(2)$ by 4^3 were studied by Alperin [5] and Griess [53]. Results on the cohomology of the Higman–Sims group can be found in [1] and [65].

Proof of Theorem 3.1. Let $g \in O'N$ be any element with $o(g) \neq 16$. Then the difference of Rademacher sums

$$-R^{[-4]}_{\frac{3}{2},4o(g)}(\tau) + 2R^{[0]}_{\frac{3}{2},4o(g)}(\tau)$$

of level $4o(g)$ is a mock modular form with the correct principal part at infinity and vanishes at all other cusps by Theorem 2.5. If $o(g)$ is even, then we know from Proposition 2.8 that this function is in the plus space. If on the other hand, $o(g)$ is odd, then we use the projection operator $|\,\mathrm{pr}$ to map it into the plus space, which by Lemma 2.6 introduces an additional pole of order $\frac{1}{4}$ at the cusp $\frac{1}{o(g)}$. This establishes the existence of a function

$$\widetilde{F}_{[g]}(\tau) := -R^{[-4],+}_{\frac{3}{2},4o(g)}(\tau) + 2R^{[0],+}_{\frac{3}{2},4o(g)}(\tau)$$

satisfying properties (a) and (b) in Theorem 3.1 (1) for $o(g) \neq 16$. To achieve this much for $o(g) = 16$ we use

$$\widetilde{F}_{[g]}(\tau) := -(2R^{[-4],+}_{\frac{3}{2},128}(\tau) - R^{[-4],+}_{\frac{3}{2},64}(\tau)) + 2(2R^{[0],+}_{\frac{3}{2},128}(\tau) - R^{[0],+}_{\frac{3}{2},64}(\tau))$$

since $\rho_{[g]}$ in this case is trivial on $\Gamma_0(128)$, and -1 on $\Gamma_0(64) \setminus \Gamma_0(128)$.

By Lemma 2.4 we see that the above properties determine a mock modular form uniquely up to cusp forms. Unless $o(g) \in \{11, 14, 15, 16, 19, 28, 31\}$ there are no cusp forms of weight $3/2$ in the plus spaces with the required characters, so one checks directly that in all those cases condition (c) is satisfied. In the remaining cases, condition (c) uniquely determines the contribution from cusp forms, because, as one can check using standard computer algebra systems (the authors used MAGMA [13] and PARI [77]), any weight $3/2$ cusp form of one of the given levels in the plus space with the relevant character is uniquely determined by the coefficients of q^3, q^4, and q^7.

We now show that the functions $F_{[g]}$ are actually all weakly holomorphic instead of just mock modular. First suppose that $o(g)$ is odd or $2||o(g)$. Then, because in those cases $o(g)$ is square-free, the shadow of $F_{[g]}(\tau)$ must be a multiple of

$$\vartheta(\tau) := \sum_{n\in\mathbb{Z}} q^{n^2},$$

which follows from the Serre-Stark basis theorem [83]. We compute Bruinier–Funke pairings (see Proposition 3.5 in [18]) and find that

$$\{\widehat{R^{[-4],+}_{\frac{3}{2},4o(g)}}(\tau),\vartheta(\tau)\}=2c \qquad \text{and} \qquad \{\widehat{R^{[0],+}_{\frac{3}{2},4o(g)}}(\tau),\vartheta(\tau)\}=c,$$

where c is some constant. This shows that the shadow of the mock modular form $F_{[g]}(\tau)=-R^{[-4],+}_{\frac{3}{2},4o(g)}(\tau)+2R^{[0],+}_{\frac{3}{2},4o(g)}(\tau)$ is 0, whence it is indeed a weakly holomorphic modular form.

If $o(g)$ is divisible by 4 or 8, but not 16, the space of possible shadows is *a priori* 2-dimensional, generated by $\vartheta(\tau)$ and $\vartheta(4\tau)$, but Proposition 2.8 and the fact that the shadow of a Rademacher sum is again a Rademacher sum show that the shadow's Fourier coefficients must be supported on exponents divisible by 4. So in fact, only multiples of $\vartheta(4\tau)$ can occur as shadows and the same computation as above shows the claim in these cases. For $o(g)=16$ the space of possible shadows is a priori one-dimensional, spanned by $\vartheta(\tau)-\vartheta(4\tau)$, but this function is supported on odd exponents so is also ruled out by the Rademacher sum construction.

It remains to show that the coefficients of the $F_{[g]}$ are all rational integers. This follows by checking finitely many coefficients and applying Sturm's Theorem [87], in a manner directly similar to the proof of Proposition 3.2 in [54], for example. As we will also see in Section 4.1, a possible bound up to which coefficients need to be checked to verify the claim is 225. ☐

4. Proof of Theorem 1.1

Here we prove that the weakly holomorphic modular forms given in Theorem 3.1 are McKay–Thompson series for the infinite-dimensional $O'N$-module W. We begin by stating a refined form of Theorem 1.1.

Theorem 4.1. *There is an infinite-dimensional graded virtual O'N-module*

$$W=\bigoplus_{\substack{m=3\\ m\equiv 0,3 \pmod 4}}^{\infty} W_m$$

such that we have

$$\mathrm{tr}(g|W_m)=a_{[g]}(m),$$

for all m. Moreover, W_m is an honest O'N-module for $m\notin\{7,8,12,16\}$ (see Tables B.1 to B.3).

We break down the proof of this theorem into separate pieces. Using the Schur orthogonality relations on the irreducible representations of $O'N$ we construct weakly holomorphic modular forms of weight $3/2$ whose coefficients are the multiplicities of the irreducible components if and only if W exists. Then the proof of Theorem 4.1 boils down to proving that these multiplicities are integral for all m, and non-negative for $m\notin\{7,8,12,16\}$. In Section 4.1 we establish integrality, and in Section 4.2 we establish the claim on non-negativity.

4.1. **Integrality of Multiplicities.** For every prime $p|\#O'N$ we find linear congruences among the alleged McKay–Thompson series. Here, we prove these, but first we note that their truth implies the following systematic congruences.

Theorem 4.2. *Let $g_j \in O'N$ of order d_j, $j = 1, 2$, with $d_2 = p^c \cdot d_1$ for some prime number p and $c \geq 1$. Then we have the congruence*

$$F_{[g_1]} \equiv F_{[g_2]} \pmod{p}.$$

In Appendix C, we list these congruences, which sometimes hold with higher prime power moduli than stated in Theorem 4.2. Assuming their correctness for the moment, we can show integrality just as described in [54]. For the convenience of the reader, we recall the method briefly.

Let $\mathbf{C} \in \mathbb{Z}^{30\times\infty}$ denote the matrix formed by the coefficients of the functions $F_{[g]}(\tau)$ for each of the 30 conjugacy classes of $O'N$ (in practice one uses a $30 \times B$ matrix for some large B). Further denote by $\mathbf{X} \in \overline{\mathbb{Q}}^{30\times 30}$ the matrix whose rows are indexed by irreducible characters and whose columns are indexed by conjugacy classes of $O'N$, with

$$\mathbf{X}_{\chi,[g]} := \frac{\overline{\chi(g)}}{\#C(g)},$$

where $C(g)$ denotes the centralizer of $g \in O'N$. By the first Schur orthogonality relation we see that the matrix

$$\mathbf{m} := \mathbf{XC}$$

gives the multiplicities of each irreducible representation in the alleged virtual representation in Theorem 4.1. Since there are repetitions among the rows of $\mathbf{C}$, because the functions $F_{[g]}(\tau)$ depend only on the order of elements in $[g]$, it does not have full rank, but by just deleting the repetitions it does turn out to have full rank, which is 18. Let $\mathbf{N}^* \in \mathbb{Z}^{18\times 30}$ denote the matrix performing this operation and let $\mathbf{N} \in \mathbb{Z}^{30\times 18}$ be the matrix that undoes it, so that

$$\mathbf{m} = \mathbf{XNN}^*\mathbf{C}.$$

Now for each prime $p|\#O'N$, we can reduce the matrix $\mathbf{N}^*\mathbf{C}$ according to the aforementioned congruences as in [54] by left-multiplying by a matrix $\mathbf{M}_p \in \mathbb{Q}^{18\times 18}$, which may be seen to have full rank. Hence we get

$$\mathbf{m} = (\mathbf{XNM}_p^{-1}) \cdot (\mathbf{M}_p\mathbf{N}^*\mathbf{C}).$$

The congruences in Appendix C ensure that the matrix $\mathbf{M}_p\mathbf{N}^*\mathbf{C} \in \mathbb{Q}^{18\times\infty}$ has all integer entries and one can check directly that the matrix $\mathbf{XNM}_p^{-1} \in \overline{\mathbb{Q}}^{30\times 18}$ has p-integral (rational) entries for every p. This shows that $\mathbf{m}$ has p-integral entries as well for each $p|\#O'N$, hence its entries must be integers, as claimed.

It remains to show the congruences. Since by Theorem 3.1 all the functions $F_{[g]}$ are weakly holomorphic modular forms, we can prove all the congruences with standard techniques from the theory of modular forms. For example, we may multiply each of the congruences by the unique cusp form g in $S^+_{\frac{25}{2}}(\Gamma_0(4))$ such that $g(\tau) = q^4 + O(q^5)$ (which has integral coefficients), thereby reducing the problem to congruences among holomorphic modular forms of weight

14. These can be checked in all cases using the Sturm bound [87], which is at most 225 in all cases.

4.2. **Positivity of Multiplicities.** Denote by $\mathrm{mult}_j(n)$ the multiplicity of the irreducible character χ_j of $O'N$ in the virtual module W_n as in Theorem 4.1, whose associated generalized character is given by the coefficients $a_{[g]}(n)$, cf. Theorem 3.1. Then the Schur orthogonality relations and the triangle inequality tell us that

$$\begin{aligned}\mathrm{mult}_j(n) &= \sum_{[g]\subseteq O'N} \frac{1}{\#C(g)} a_{[g]}(n)\overline{\chi_j(g)} \\ &\geq \frac{|a_1(n)|}{\#O'N}\chi_j(1) - \sum_{[g]\neq 1A} \frac{|a_{[g]}(n)|}{\#C(g)}|\chi_j(g)|,\end{aligned} \tag{4.1}$$

where the summations run over conjugacy classes of $O'N$. Hence in order to show the eventual positivity of all $\mathrm{mult}_j(n)$, we want to establish explicit lower bounds on $a_{1A}(n)$, and upper bounds on $a_{[g]}(n)$ for $g \neq 1$. Recall that

$$F_{[g]}(\tau) = -q^{-4} + 2 + \sum_{n=1}^{\infty} a_{[g]}(n)q^n = -R^{[-4],+}_{\frac{3}{2},4o(g)}(\tau) + 2R^{[0],+}_{\frac{3}{2},4o(g)}(\tau) + \text{cusp form}$$

for $o(g) \neq 16$, and

$$F_{[g]}(\tau) = -(2R^{[-4],+}_{\frac{3}{2},128}(\tau) - R^{[-4],+}_{\frac{3}{2},64}(\tau)) + 2(2R^{[0],+}_{\frac{3}{2},128}(\tau) - R^{[0],+}_{\frac{3}{2},64}(\tau))$$

for $o(g) = 16$. (For $o(g) = 16$ we have $\dim S^{+}_{\frac{3}{2}}(\Gamma_0(4o(g)), \rho_{[g]}) = 1$ but condition (c) in Theorem 3.1 rules out any cuspidal contribution to $F_{[g]}$. Cf. Appendix D.)

We bound each of the components individually, following the strategy already employed in [37, 49, 54], which we sketch briefly for the convenience of the reader. Note however that in the cited papers, only the coefficients of one Rademacher sum had to be considered, since the corrections there were known to come from weight $\frac{1}{2}$ modular forms, whose coefficients are bounded. In our case the corrections can grow with n.

Since the computations necessary to bound the contribution coming from the Rademacher sum $R^{[-4]}_{\frac{3}{2},4N}$, which is obviously going to be the dominant part, have been carried out in detail in [49, 54], we omit them here. The idea is to use the known formula for the coefficients of the Rademacher sum in terms of infinite sums of Kloosterman sums weighted by I-Bessel functions, see Section 2. One then splits this sum into three parts, a dominant part, an absolutely convergent remainder term and a value of a *Selberg–Kloosterman zeta function*, the first two of which may be bounded by elementary means, and for the third, one uses Proposition 4.1 in [54] (which we note is directly applicable to our situation).

4.2.1. *Bounding Coefficients of Rademacher Sums.* From Proposition 5.2 below, we see that the $\mu = 0$ Rademacher sum can be explicitly given in terms of generating functions of generalized Hurwitz class numbers $H^{(N)}(n)$ (see Section 5 for the definition). While strong bounds for class numbers are known (see for instance Chapter 23 in [63] and the references therein), they are usually not explicit. For our purposes, crude bounds on class numbers suffice.

Proposition 4.3. *For every $N \in \mathbb{N}$, $-D \leq -5$ a negative discriminant and $\varepsilon > 0$ we have*

$$H^{(N)}(D) \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] c_\varepsilon D^\varepsilon \cdot \frac{\sqrt{D}}{2\pi}\left(1 + \frac{1}{2}\log D\right),$$

where we can choose

$$c_\varepsilon = \prod_{p < e^{\frac{1}{2\varepsilon}}} (2\varepsilon p^{1/\log p - 2\varepsilon} \log p)^{-1}.$$

Proof. First we note that we trivially have the bound $H^{(N)}(D) \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] H^{(1)}(D)$ by definition. Now suppose for the moment that D is a fundamental discriminant. Then Dirichlet's class number formula gives

$$H(D) = \frac{\sqrt{D}}{2\pi} \cdot L\left(1, \left(\frac{-D}{\bullet}\right)\right).$$

Theorem 13.3 of Chapter 12 in [61] tells us that for $D \geq 5$ we have the upper bound

$$L\left(1, \left(\frac{-D}{\bullet}\right)\right) < 1 + \frac{1}{2}\log D.$$

By [97, pp. 73f.], Dirichlet's formula is also valid for non-fundamental discriminants if only primitive forms are counted, so that we get the bound

$$H(D) \leq \tau_\square(D) \frac{\sqrt{D}}{2\pi}\left(1 + \frac{1}{2}\log D\right),$$

where $\tau_\square(n)$ denotes the number of square divisors of n. Considering the prime factorisation of D, it is elementary to see that $\tau_\square(D) \leq c_\varepsilon D^\varepsilon$ for any $\varepsilon > 0$ and c_ε as claimed. ☐

This result together with Proposition 5.2 gives a sufficient and explicit bound for the coefficients of the Rademacher sum $R^{[0],+}_{\frac{3}{2},4N}$. For the actual computations we choose $\varepsilon = \frac{1}{8}$, which yields $c_\varepsilon \approx 10.6766$.

4.2.2. *Bounding Coefficients of Cusp Forms.* For $g \in O'N$ with

$$o(g) \in \{11, 14, 15, 19, 28, 31\},$$

there are non-trivial cusp forms in $S^+_{\frac{3}{2}}(\Gamma_0(4o(g)))$ contributing to our modular forms $F_{[g]}$, see Appendix D. According to the Ramanujan–Petersson conjecture, the coefficients of these cusp forms should grow like $O(n^{\frac{1}{4}+\varepsilon})$ (for n square-free). Unconditional bounds (again for square-free n) have been obtained by Iwaniec [62] for weights $\geq 5/2$ and Duke [32] for weight 3/2 (see also [33]). These bounds have one main disadvantage for our purposes, namely that the constants involved in them are not explicit or not computable. Here, we outline how to give completely explicit and computable, but very crude, estimates for the cusp form coefficients in question.

Let $P^{[m]}_{4N}$ denote the cuspidal Poincaré series of weight 3/2 characterized by the *Petersson coefficient formula*,

$$\langle f, P^{[m]}_{4N} \rangle = \frac{b_f(m)}{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(4N)]\sqrt{4m}} \quad \forall\ f(\tau) = \sum_{n=1}^{\infty} b_f(n) q^n \in S^+_{\frac{3}{2}}(\Gamma_0(4N)), \tag{4.2}$$

where the Petersson inner product on $S^+_{\frac{3}{2}}(\Gamma_0(4N))$ is defined by the usual double integral

$$\langle f_1, f_2\rangle = \frac{1}{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(4N)]}\int_{\Gamma_0(4N)\backslash\mathfrak{H}} f_1(\tau)\overline{f_2(\tau)}y^{\frac{3}{2}}\frac{du\,dv}{v^2}.$$

The Fourier coefficients of these Poincaré series are given in terms of infinite sums of Kloosterman sums times J-Bessel functions (see Proposition 4 in [66]), and essentially the same computation used to bound the coefficients of the Rademacher sums $R^{[-4]}_{\frac{3}{2},4N}$ can be used here as well. It is then only necessary to express the cusp forms $\mathscr{G}^{(o(g))}$ (see again Appendix D) in terms of these Poincaré series, which is particularly easy in the cases where $o(g) \neq 31$ is odd, since in those cases, the space $S^+_{\frac{3}{2}}(\Gamma_0(4o(g)))$ is one-dimensional and $\mathscr{G}^{(o(g))}$ is a newform. Hence we have $\langle \mathscr{G}^{(o(g))}, P^{[m]}_{4N}\rangle = \beta\langle \mathscr{G}^{(o(g))}, \mathscr{G}^{(o(g))}\rangle$, where we choose m to be the order of $\mathscr{G}^{(o(g))}$ at ∞. It therefore remains to compute the Petersson norm of the newform $\mathscr{G}^{(o(g))}$. This can be done by means of the following result due to Kohnen, which is an explicit version of Waldspurger's theorem (see Corollary 1 in [66]).

Proposition 4.4. *Let $N \in \mathbb{N}$ be odd and square-free, $f \in S^+_{k+\frac{1}{2}}(\Gamma_0(4N))$ be a newform and $F \in S_{2k}(\Gamma_0(N))$ the image of f under the Shimura correspondence. For a prime $\ell | N$, let w_ℓ be the eigenvalue of F under the Atkin–Lehner involution W_ℓ and choose a fundamental discriminant D with $(-1)^k D > 0$ and $\left(\frac{D}{\ell}\right) = w_\ell$ for all ℓ. Then we have*

$$\langle f, f\rangle = \frac{\langle F, F\rangle \pi^k}{2^{\omega(N)}(k-1)!|D|^{k-\frac{1}{2}}L(F, D; k)} \cdot |b_f(|D|)|^2,$$

where $L(F, D; s)$ denotes the twist of the newform F by the quadratic character $\left(\frac{D}{\bullet}\right)$ and $\omega(N)$ denotes the number of distinct prime divisors of N.

Since the twisted L-series has a functional equation of the usual type, there are efficient methods to compute its values numerically. (The authors used the built-in intrinsics of Magma [13].) Computing the Petersson norm of F is also possible to high accuracy, e.g. by using the well-known relationship (cf. [30, 98])

$$\langle F, F\rangle = \frac{\mathrm{vol}(E)}{4\pi^2}\deg(\varphi_E),$$

where F is the newform associated to the elliptic curve $E/\mathbb{Q}$, we denote the covolume of the period lattice of E by $\mathrm{vol}(E)$, and use φ_E for the modular parametrization of E. (Every elliptic curve E we consider in this paper has $\deg(\varphi_E) = 1$.) Alternatively, the Petersson norm of F is computed for N prime by Theorem 2 in [98].

Remark. Kohnen's result Proposition 4.4 has been extended to many situations, e.g. by Ueda and his collaborators [91, 92] to certain even levels and forms not in the plus-space (see in particular Corollary 1 in [69]). So the above reasoning carries over to $o(g) \in \{14, 28\}$, by noting that $\mathscr{G}^{(14)}$ and $\mathscr{G}^{(28)}$ both arise from the unique normalized cusp form in $S_{\frac{3}{2}}(\Gamma_0(28))$ (not in the plus space). The former may be obtained by applying sieve operators, the latter by applying the V_4-operator.

Remark. For $o(g) = 31$, the above reasoning only needs to be modified to take into account that $\mathscr{G}^{(31)}$ is not a Hecke eigenform, but its decomposition into newforms is given in Appendix D. Using the fact that these newforms are orthogonal, the only difference becomes that one needs to take into account two Poincaré series instead of one.

Putting the estimates for the Rademacher sums $R^{[-4]}_{\frac{3}{2},4N}$, $R^{[0]}_{\frac{3}{2},4N}$, and the occuring cusp forms together and plugging them all into (4.1), one finds that the multiplicities are nonnegative as soon as $n \geq 113$ (the worst case occurs for the character χ_1). Inspecting the remaining coefficients by computer then completes the proof of Theorem 4.1.

5. TRACES OF SINGULAR MODULI

In this section, we discuss and recall some basic notation and facts about traces of singular moduli. Their study originates in seminal work by Zagier [99], and has since been an important subject in number theory (cf. for instance [8, 16, 19, 70], just to name a few). They appeared in connection with moonshine for the Thompson group in [58].

5.1. **Genus Zero Levels.** It is well-known that for $N \in \{1,2,3,4,5,6,7,8,10,12,16\}$ the modular curve $X_0(N)$ has genus 0, so that in those cases, there is a Hauptmodul $J^{(N)}$. These Hauptmoduln are given explicitly in Table 5.1 in terms of the Dedekind eta function $\eta(\tau) := q^{\frac{1}{24}}\prod_{n>0}(1-q^n)$ and the Eisenstein series $E_4(\tau) := 1 + 240\sum_{n>0} n^3q^n(1-q^n)^{-1}$.

N	$J^{(N)}(\tau)$
1	$\frac{E_4(\tau)^3}{\eta(\tau)^{24}} - 744$
2	$\frac{\eta(\tau)^{24}}{\eta(2\tau)^{24}} + 24$
3	$\frac{\eta(\tau)^{12}}{\eta(3\tau)^{12}} + 12$
4	$\frac{\eta(\tau)^{8}}{\eta(4\tau)^{8}} + 8$
5	$\frac{\eta(\tau)^{6}}{\eta(5\tau)^{6}} + 6$
6	$\frac{\eta(\tau)^{5}\eta(3\tau)}{\eta(2\tau)\eta(6\tau)^{5}} + 5$
7	$\frac{\eta(\tau)^{4}}{\eta(7\tau)^{4}} + 4$
8	$\frac{\eta(\tau)^{4}\eta(4\tau)^2}{\eta(2\tau)^2\eta(8\tau)^{4}} + 4$
10	$\frac{\eta(\tau)^{3}\eta(5\tau)}{\eta(2\tau)\eta(10\tau)^{3}} + 3$
12	$\frac{\eta(\tau)^{3}\eta(4\tau)\eta(6\tau)^2}{\eta(2\tau)^2\eta(3\tau)\eta(12\tau)^{3}} + 3$
16	$\frac{\eta(\tau)^{2}\eta(8\tau)}{\eta(2\tau)\eta(16\tau)^{2}} + 2$

TABLE 5.1. Hauptmoduln for some $\Gamma_0(N)$

To make use of these Hauptmoduln we require some notation. Denote by $\mathcal{Q}^{(N)}_{-D}$ the set of positive definite quadratic forms $Q = ax^2 + bxy + cy^2 =: [a,b,c]$ of discriminant $-D = b^2 - 4ac < 0$ such that $N|a$. It is well-known that $\Gamma_0(N)$ acts on $\mathcal{Q}^{(N)}_{-D}$ with finitely many orbits, which correspond to the so-called *Heegner points* on the modular curve $X_0(N)$. For $Q = [a,b,c] \in \mathcal{Q}^{(N)}_{-D}$, we denote by $\tau_Q := \frac{-b+i\sqrt{D}}{2a}$ the unique root of $Q(x,1)$ in $\mathfrak{H}$. For a function $f : \mathfrak{H} \to \mathbb{C}$ invariant under the action of $\Gamma_0(N)$ we then define the trace function

$$\mathrm{Tr}^{(N)}_D(f) := \sum_{Q \in \mathcal{Q}^{(N)}_{-D}/\Gamma_0(N)} \frac{f(\tau_Q)}{\omega^{(N)}(Q)}, \tag{5.1}$$

where $\omega^{(N)}(Q) = \frac{1}{2} \cdot \# \operatorname{Stab}_{\Gamma_0(N)}(Q)$. Further let

$$\mathscr{H}^{(N)}(\tau) := -\frac{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)]}{12} + \sum_{\substack{D>0 \\ D\equiv 0,3 \pmod 4}} H^{(N)}(D) q^D$$

denote the generating function of the *(generalized) Hurwitz class numbers of level N* which are defined as $H^{(N)}(D) := \operatorname{Tr}_D^{(N)}(1)$. The special case of $N = 1$ yields the classical Hurwitz class numbers $H^{(1)}(D) := H(D)$.

It is a straightforward consequence of Theorem 1.2 in [70], analogous to Theorem 1.2 in [8], that we can describe the Fourier coefficients of the Rademacher sums $R^{[-4],+}_{\frac{3}{2},4N}$ as traces of the Hauptmoduln in Table 5.1.

Proposition 5.1. *Let $N \in \mathbb{N}$ such that $X_0(N)$ has genus 0 and*

$$\operatorname{Tr}_4^{(N)}(D) := \frac{1}{2}\left(\operatorname{Tr}_D^{(N)}(J_2^{(N)}) - \operatorname{Tr}_D^{(N/d)}(J^{(N/d)})\right), \tag{5.2}$$

where $J_2^{(N)} = q^{-2} + O(q)$ is the unique modular function for $\Gamma_0(N)$ with this Fourier expansion at infinity and no poles anywhere else and $d := \gcd(N, 2)$. Then we have

$$\mathscr{T}^{(N)}(\tau) := -q^{-4} + \sum_{\substack{D>0 \\ D\equiv 0,3 \pmod 4}} \operatorname{Tr}_4^{(N)}(D) q^D = -R^{[-4],+}_{\frac{3}{2},4N}(\tau) - \frac{c_2}{2}\mathscr{H}^{(N)}(\tau) + \frac{c_1}{2}\mathscr{H}^{(N/d)}(\tau) \tag{5.3}$$

for certain rational numbers c_1 and c_2. In particular, the function $\mathscr{T}^{(N)}$ has integer Fourier coefficients.

Remark. It should be pointed out that Theorem 1.2 in [70] is only stated for odd levels, although the proof goes through for even levels as well.

Remark. The rational numbers c_2 and c_1 in Proposition 5.1 are the constant terms of the weight 0 Rademacher sums $R^{[-2]}_{0,N}$ and $R^{[-1]}_{0,N/d}$, respecitvely. For a proof of the rationality of these numbers see Lemma 3.2 in [8].

For the Rademacher sum $R^{[0],+}_{\frac{3}{2},4N}$ we get the following.

Proposition 5.2. *For $N \in \mathbb{N}$ we have*

$$R^{[0],+}_{\frac{3}{2},4N} = -\frac{12}{\varphi(N)} \sum_{d|N} \frac{d}{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(d)]} \mu\left(\frac{N}{d}\right) \mathscr{H}^{(d)}$$

where μ and φ denote the Möbius function and Euler's totient function, respectively.

Proof. This follows from a straightforward modification of the proof of Theorem 1.2 in [70]. □

Note that the above Proposition 5.2 is indeed valid for all N, not just those such that $X_0(N)$ has genus 0.

Putting Propositions 5.1 and 5.2 together we obtain explicit descriptions of the functions $F_{[g]}$ in terms of singular moduli for $o(g) \in \{1, 2, 3, 4, 5, 6, 7, 8, 10, 12\}$. These are given in Appendix D.

5.2. **Positive Genus Levels.** In the remaining cases, i.e. where

$$o(g) \in \{11, 14, 15, 16, 19, 20, 28, 31\},$$

our $F_{[g]}$ involve (cf. the proof of Theorem 3.1) Rademacher sums $R^{[-4],+}_{\frac{3}{2},4N}$ where N is such that $X_0(N)$ has positive genus. So there is no notion of a Hauptmodul there. However, it is known that for all these levels, the modular curve $X_0^+(N)$, being the quotient of $X_0(N)$ by all Atkin–Lehner involutions, does have genus 0 (see e.g. [43]). So there exists a Hauptmodul $J^{(N,+)}(\tau)$ for the corresponding group $\Gamma_0^+(N)$. See Table 5.2 for these. There, $E_2(\tau) := 1 - 24\sum_{n>0} nq^n(1-q^n)^{-1}$ is the quasimodular Eisenstein series, $f_{19} = q - 2q^3 + O(q^4)$ denotes the weight 2 newform associated to the elliptic curve

$$E_{19} : y^2 + y = x^3 + x^2 - 9x - 15$$

([68, Elliptic Curve 19.a2]), and $f_{31} = q + \frac{1+\sqrt{5}}{2}q^2 + O(q^3)$ denotes the unique newform in $S_2(\Gamma_0(31))$ up to Galois conjugation (which is denoted by an exponent σ).

N	11	14
$J^{(N,+)}(\tau)$	$-\frac{E_2(\tau)-11E_2(11\tau)}{10\eta(\tau)^2\eta(11\tau)^2} - \frac{22}{5}$	$-\frac{E_2(\tau)+2E_2(2\tau)-7E_2(7\tau)-14E_2(14\tau)}{18\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau)} - \frac{7}{3}$

N	15	16
$J^{(N,+)}(\tau)$	$-\frac{E_2(\tau)+3E_2(3\tau)-5E_2(5\tau)-15E_2(15\tau)}{16\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau)} - \frac{5}{2}$	$\frac{\eta(2\tau)^6\eta(8\tau)^6}{\eta(\tau)^4\eta(4\tau)^4\eta(16\tau)^4} - 4$

N	19	20	28
$J^{(N,+)}(\tau)$	$-\frac{E_2(\tau)-19E_2(19\tau)}{18f_{19}(\tau)} - \frac{4}{3}$	$\frac{\eta(2\tau)^8\eta(10\tau)^8}{\eta(\tau)^4\eta(4\tau)^4\eta(5\tau)^4\eta(20\tau)^4} - 4$	$\frac{\eta(2\tau)^6\eta(14\tau)^6}{\eta(\tau)^3\eta(4\tau)^3\eta(7\tau)^3\eta(28\tau)^3} - 3$

N	31	32
$J^{(N,+)}(\tau)$	$\frac{\sqrt{5}(f_{31}(\tau)+f_{31}^{\sigma}(\tau))}{2(f_{31}(\tau)-f_{31}^{\sigma}(\tau))} - \frac{5}{2}$	$\frac{\eta(2\tau)^3\eta(16\tau)^3}{\eta(\tau)^2\eta(4\tau)\eta(8\tau)\eta(32\tau)^2} - 2$

TABLE 5.2. Hauptmoduln for some $\Gamma_0^+(N)$

Armed with these Hauptmoduln we can now express the Fourier coefficients of all the remaining Rademacher sums $R^{[-4],+}_{\frac{3}{2},4N}$ in terms of singular moduli of holomorphic modular functions and class numbers.

Proposition 5.3. *Let $N \in \mathbb{N}$ such that $X_0^+(N)$ has genus 0 and define*

$$\mathrm{Tr}_4^{(N,+)}(D) := \frac{1}{2}\left(\frac{1}{2^{\omega(N)}}\,\mathrm{Tr}_D^{(N)}\left(J_2^{(N,+)}\right) - \frac{1}{2^{\omega(N/d)}}\,\mathrm{Tr}_D^{(N/d)}\left(J^{(N/d,+)}\right)\right) \tag{5.4}$$

where $J_2^{(N,+)} = q^{-2} + O(q)$ *is the unique modular function for* $\Gamma_0^+(N)$ *with this Fourier expansion at infinity and no poles anywhere else and* $d := \gcd(N,2)$*. Then we have*

(5.5)

$$\mathscr{T}^{(N,+)}(\tau) := -q^{-4} + \sum_{\substack{D>0\\ D\equiv 0,3 \pmod 4}} \mathrm{Tr}_4^{(N,+)}(D)q^D = -R^{[-4],+}_{\frac{3}{2},4N}(\tau) - \frac{c_2}{2}\mathscr{H}^{(N)}(\tau) + \frac{c_1}{2}\mathscr{H}^{(N/d)}(\tau)$$

for some rational numbers c_1 *and* c_2*, where* $\omega(N)$ *denotes the number of distinct prime factors of* N*.*

Proof. Proposition 5.1 turns out to be valid for all N, if one replaces the Hauptmodul $J^{(N)}$ by the completed Rademacher sum $\widehat{R^{[-1]}_{0,N}}$, normalized so that its constant term is 0 and $J_2^{(N)}$ by $\widehat{R^{[-2]}_{0,N}}$ with the same normalization, which is the original formulation in [70]. Note that these Rademacher sums coincide with the Hauptmoduln where applicable. Now we consider for $N'||N$, i.e. $\gcd(N', N/N') = 1$, the Atkin–Lehner involution $W_{N'}$. These involutions map the set $\mathcal{Q}^{(N)}_{-D}/\Gamma_0(N)$ bijectively to itself (see e.g. Section 1 of [56]), so for any $\Gamma_0(N)$-invariant function f we have that $\mathrm{Tr}_D^{(N)}(f) = \mathrm{Tr}_D^{(N)}(f|W_{N'})$. Since there are exactly $2^{\omega(N)}$ Atkin–Lehner involutions of level N this means that

$$\mathrm{Tr}_D^{(N)}(f) = \frac{1}{2^{\omega(N)}}\mathrm{Tr}_D^{(N)}(\tilde{f})$$

where $\tilde{f} := \sum_{N'||N} f|W_{N'}$. The function $\tilde{f}$ is clearly $\Gamma_0^+(N)$-invariant. By checking the polar parts we conclude that if f is $\widehat{R^{[-1]}_{0,N}}$ or $\widehat{R^{[-2]}_{0,N}}$ then $\tilde{f}$ has to coincide with $J^{(N,+)}$ or $J_2^{(N,+)}$, respectively, up to a rational additive constant. This proves the result. ☐

We now put Propositions 5.2 and 5.3 together to obtain explicit descriptions of the functions $F_{[g]}$ in terms of singular moduli for $o(g) \in \{11, 14, 15, 19, 28, 31\}$. For $o(g) = 16$ we use Proposition 5.1 as well since $F_{[g]}$ involves Rademacher sums $R^{[\mu],+}_{\frac{3}{2},4N}$ for both $N = 16$ and $N = 32$ in this case. The resulting expressions are given in Appendix D.

6. NUMBER THEORETIC APPLICATIONS

In this section we prove the arithmetic applications of O'Nan moonshine given in Theorems 1.2 to 1.4. All these proofs rely on the following easy observation.

Lemma 6.1. *Let* $N > 1$ *be an integers and* $-D < 0$ *a discriminant which is not a square in* $\mathbb{Z}/N\mathbb{Z}$*. Then the set* $\mathcal{Q}^{(N)}_{-D}$ *is empty. In particular, we have that* $\mathrm{Tr}_{-D}^{(N)}(f) = H^{(N)}(D) = 0$ *for any* $\Gamma_0(N)$*-invariant function* f*.*

Proof. A quadratic form $[a,b,c] \in \mathcal{Q}^{(N)}_{-D}$ satisfies $-D = b^2 - 4ac$ and $N|a$, hence if $-D$ is not a square modulo N, there cannot be any such forms. ☐

6.1. **Proof of Theorem 1.2.**

Suppose first that $p \in \{5, 7\}$ and let D be as in Theorem 1.2. The congruences in Appendix C together with the identities in Appendix D imply the congruence

$$\dim(W_D) \equiv \mathrm{tr}(g_p|W_D) \equiv \mathrm{Tr}_4^{(p)}(D) - 24H(D) + \alpha_p H^{(p)}(D) \pmod p$$

for some integer α_p. By Lemma 6.1, the terms $\mathrm{Tr}_4^{(p)}(D)$ and $H^{(p)}(D)$ vanish for D as required, proving the result. For $p = 3$, one replaces the modulus above by 3^2, making the congruence non-trivial.

For $p = 2$, we note that there is a congruence between $\dim W_D$ and $\mathrm{tr}(g_2|W_D)$ modulo 2^{11} by Appendix C. As one easily sees through a Sturm bound argument, we also have

$$\dim W_D \equiv \mathrm{tr}(g_2|W_D) \equiv 0 \pmod{16}$$

for $D \equiv 4, 8 \pmod{16}$, which is the case in particular when $-D < 8$ is an even fundamental discriminant. The fact that for these D the class number is even can be seen in various ways, for example by noting that by a famous theorem of Gauss and Hermite we have that $24H(D) = 2r_3(D/4)$, where $r_3(n)$ is the number of representations of n as the sum of three squares. Since $-D$ is fundamental, it follows that $D/4$ is square-free and hence is not the sum of three or just two equal squares. Through an easy case-by-case analysis one then finds that $r_3(D/4)$ is always divisible by 8. Alternatively, one could also show the modular forms congruence

$$\sum_{n\equiv 1,2 \pmod{4}} r_3(n)q^n \equiv 6\sum_{n=0}^{\infty} q^{(2n+1)^2} + 4\sum_{n=0}^{\infty} q^{2(2n+1)^2} \pmod{8}.$$

This completes the proof.

6.2. **Preliminaries on Elliptic Curves.** The proofs of Theorems 1.3 and 1.4 require a little preparation which we provide in this section.

One of the most important open problems in the theory of elliptic curves is the Birch and Swinnerton-Dyer Conjecture.

Conjecture 6.2. *Let $E/\mathbb{Q}$ be an elliptic curve. Then we have that*

$$\frac{L^{(r)}(E,1)}{r!\Omega_E} = \frac{\#\text{Ш}(E)\cdot \mathrm{Reg}(E)\prod_\ell c_\ell(E)}{(\#E(\mathbb{Q})_{tors})^2}, \tag{6.1}$$

where r denotes the order of vanishing of $L(E,s)$ at $s = 1$, which equals the Mordell–Weil rank of E, Ω_E is the real period of E, $\#\text{Ш}(E)$ and $\mathrm{Reg}(E)$ denote the order of the Tate-Shafarevich group and the regulator of E, respectively, the $c_\ell(E)$ for prime ℓ are the Tamagawa numbers of E, and $\#E(\mathbb{Q})_{tors}$ signifies the order of the torsion subgroup of the $\mathbb{Q}$-rational points of E.

The weak Birch and Swinnerton-Dyer conjecture—that the order of vanishing of $L(E,s)$ at $s = 1$ equals the rank of E—was established for curves of ranks 0 and 1 through work of Gross–Zagier [57] and Kolyvagin [67]. More recently, Bhargava–Shankar [9] proved, using Kolyvagin's theorem and the proof of the Iwasawa main conjectures for GL_2 by Skinner–Urban [86] (among other deep results), that a positive proportion of all elliptic curves satisfy the weak Birch and Swinnerton-Dyer Conjecture.

It is known that the left-hand side of (6.1) is always a rational number, see for instance [3, Theorem 3.2]. The following result shows that in certain situations, a local version of Conjecture 6.2, which is going to be sufficient for our purposes, holds.

Theorem 6.3 ([85], Theorem C)**.** *Let $E/\mathbb{Q}$ be an elliptic curve and $p \geq 3$ a prime of good ordinary or multiplicative reduction. Further assume that the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-representation $E[p]$ is irreducible and that there exists a prime $p' \neq p$ at which E has multiplicative reduction and $E[p]$ ramifies. If $L(E,1) \neq 0$, then we have that*

$$\mathrm{ord}_p\left(\frac{L(E,1)}{\Omega_E}\right) = \mathrm{ord}_p\left(\#\text{Ш}(E)\prod_\ell c_\ell(E)\right).$$

If $L(E,1) = 0$, then we have $\mathrm{Sel}(E)[p] \neq \{0\}$.

We are especially interested in quadratic twists of elliptic curves. In this context, the following result by Agashe, giving the real period of such a twist, turns out to be very useful.

Lemma 6.4 ([4], Lemma 2.1)**.** *Let $E/\mathbb{Q}$ be an elliptic curve of conductor N and let $-D < 0$ be a fundamental discriminant coprime to N. Then we have that*

$$\Omega_{E(-D)} = c_E \cdot c_\infty(E(-D)) \cdot \omega_-(E)/\sqrt{D},$$

where c_E denotes the Manin constant of E, $c_\infty(E(-D))$ denotes the number of components of $E(-D)$ over $\mathbb{R}$, and $\omega_-(E)$ denotes the second period of the period lattice of E.

Remark. The famous Manin Conjecture states that $c_E = 1$.

Combining this with a theorem of Kohnen [66] (cf. Proposition 4.4), we obtain the following.

Lemma 6.5. *Let $E/\mathbb{Q}$ be an elliptic curve of odd, square-free conductor N and let $-D < 0$ be a fundamental discriminant satisfying $\left(\frac{-D}{\ell}\right) = w_\ell$, where w_ℓ denotes the eigenvalue of the newform $F_E \in S_2(\Gamma_0(N))$ associated to E and the Atkin–Lehner involution W_ℓ, $\ell|N$. Denote by D_0 the smallest such discriminant. Further let $f_E(\tau) = \sum_{n=3}^\infty b_E(n)q^n \in S^+_{\frac{3}{2}}(\Gamma_0(4N))$ be the weight $3/2$ cusp form associated to F_E under the Shintani lift. For $p \geq 3$ prime we then have that*

$$\mathrm{ord}_p\left(\frac{L(E(-D),1)}{\Omega_{E(-D)}}\right) = \mathrm{ord}_p\left(\frac{L(E(-D_0),1)}{\Omega_{E(-D_0)}}\right) + \mathrm{ord}_p\left(b_E(|D|)^2\right).$$

Proof. By combining Proposition 4.4 and Lemma 6.4, we find for the fundamental discriminants $-D < 0$ as in the lemma that

$$\frac{L(E(-D),1)}{\Omega_{E(-D)}} = \frac{\pi\langle F,F\rangle}{c_E \cdot c_\infty(E(-D))2^{\omega(N)}\langle f,f\rangle\omega_-(E)} \cdot |b_E(D)|^2. \tag{6.2}$$

We see that the only quantities in this formula depending on D are $c_\infty(E(-D))$ and $b_E(D)$. Since the former is always either 1 or 2 and p is odd, it doesn't affect the p-adic valuation at all, which proves the lemma. □

Remark. If the conductor N is even but still square-free, the same result still holds along the same lines, using the remark following Proposition 4.4. The exact formula in this case only differs from (6.2) by a power of 2, which doesn't affect the p-adic valuation.

6.3. **Proofs.** In this section, we prove Theorems 1.3 and 1.4. The proofs of both theorems are very similar in their main steps, so we combine them here.

Proof of Theorems 1.3 and 1.4. By applying the expressions for the relevant $F_{[g]}$ in terms of the traces of singular moduli, class numbers and weight 3/2 cusp forms in Appendix D, and using the congruences in Appendix C, we find that

$$\begin{aligned}
\dim(W_D) \equiv \operatorname{tr}(g_{11}|W_D) &\equiv \mathrm{Tr}_4^{(11)}(D) - 24H(D) + \alpha_{11}H^{(11)}(D) + \gamma_{11}b_{11}(D) \pmod{11},\\
\operatorname{tr}(g_2|W_D) \equiv \operatorname{tr}(g_{14}|W_D) &\equiv \mathrm{Tr}_4^{(14)}(D) + \delta_7(H(D) - \delta_7 H^{(2)}(D))\\
&\qquad + \alpha_7 H^{(7)}(D) + \beta_7 H^{(14)}(D) + \gamma_7 b_{14}(D) \pmod{7},\\
\operatorname{tr}(g_3|W_D) \equiv \operatorname{tr}(g_{15}|W_D) &\equiv \mathrm{Tr}_4^{(15)}(D) + \delta_5(H(D) - \delta_5 H^{(3)}(D))\\
&\qquad + \alpha_5 H^{(5)}(D) + \beta_5 H^{(15)}(D) + \gamma_5 b_{15}(D) \pmod{5},\\
\dim(W_D) \equiv \operatorname{tr}(g_{19}|W_D) &\equiv \mathrm{Tr}_4^{(19)}(D) - 24H(D) + \alpha_{19}H^{(19)}(D) + \gamma_{19}b_{19}(D) \pmod{19},
\end{aligned}$$

where $\delta_p = \frac{p-1}{2}$, α_p, β_p are some integers, γ_p are p-adic units, and $b_N(D)$ denotes the D^{th} coefficient of the weight 3/2 cusp form $\mathscr{G}^{(N)}$ specified in Appendix D. If $-D$ is a fundamental discriminant as specified in Theorems 1.3 and 1.4 respectively, then by Lemma 6.1, the terms $\mathrm{Tr}_4^{(N)}(D)$ as well as $H^{(p)}(D)$ and $H^{(N)}(D)$ above disappear. This shows that the class number congruences in our theorems hold if and only if p divides the coefficient $b_N(D)$, i.e. if and only if $\mathrm{ord}_p\left(\frac{L(E_N(-D),1)}{\Omega_{E_N(-D)}}\right) > 0$ by Lemma 6.5. (A Magma computation reveals that the ratio $\frac{L(E_N(-D_0),1)}{\Omega_{E_N(-D_0)}}$ for the smallest possible D_0 is in each case a p-adic unit.)

Suppose for simplicity that $L(E_N(-D),1) \neq 0$. According to the Birch and Swinnerton-Dyer Conjecture 6.2, this implies that

$$\mathrm{ord}_p(\#\text{Ш}(E_N(-D))\prod_\ell c_\ell(E(-D))) > 0,$$

so our theorems follow, conditionally on Conjecture 6.2, if the Tamagawa numbers $c_\ell(E(-D))$ are never divisible by p in our cases. To establish this, we note (cf. [84, Appendix C, Table 15.1]) that for an elliptic curve $E/\mathbb{Q}$ we have that $p|c_\ell(E)$ if and only if the reduction type of E at ℓ is I_n with $p|n$, which means that $\mathrm{ord}_\ell(\Delta(E)) = n$, where $\Delta(E)$ denotes the (minimal) discriminant of E. An inspection of Tate's algorithm for the computation of Tamagawa numbers and the well-known formulas for minimal discriminants from the Kraus–Laska algorithm reveals that in our case, because we are considering twists of elliptic curves by fundamental discriminants, all the Tamagawa numbers must be in $\{1,2,3,4\}$. The argumentation in the case $L(E(-D),1) = 0$ is similar. This completes the proof of Theorem 1.3 for $N \in \{11, 19\}$.

The truth of Theorem 1.4 does not depend on the Birch and Swinnerton-Dyer Conjecture, but rather on Skinner's Theorem 6.3. A lemma of Serre [82, §2.8, Corollaire, p. 284] shows that the Galois representations $E_{14}(-D)[7]$ and $E_{15}(-D)[5]$ are surjective and hence irreducible. Furthermore, it is immediate to check that $E_{14}(-D)$ (resp. $E_{15}(-D)$) has multiplicative reduction modulo 2 (resp. 3) and that $E_{14}(-D)[7]$ (resp. $E_{15}(-D)[5]$) ramifies there, so the conditions of Theorem 6.3 are satisfied, completing the proof of Theorem 1.4. ☐

7. Examples

Here we offer some numerical examples which illustrate the congruences described in the introduction.

7.1. **Class Number Congruences.** Here we present some class number congruences that arise from Theorem 1.2. Recall that this theorem offers congruences modulo 16, 9, 5, and 7 for certain fundamental discriminants $-D < 0$ which satisfy given congruence conditions. The three columns in Tables 7.1 to 7.4 are congruent, which illustrates the theorem.

D	$\dim W_D$	$\mathrm{tr}(g_2\|W_D)$	$-24H(D)$
20	$798588584512 \equiv 0 \pmod{16}$	$576 \equiv 0 \pmod{16}$	$-48 \equiv 0 \pmod{16}$
24	$116700...6880 \equiv 0 \pmod{16}$	$-1088 \equiv 0 \pmod{16}$	$-48 \equiv 0 \pmod{16}$
40	$905977...8912 \equiv 0 \pmod{16}$	$-10304 \equiv 0 \pmod{16}$	$-48 \equiv 0 \pmod{16}$

Table 7.1. $p = 2$

D	$\dim W_D$	$\mathrm{tr}(g_3\|W_D)$	$-24H(D)$
4	$143376 \equiv 6 \pmod{9}$	$6 \equiv 6 \pmod{9}$	$-12 \equiv 6 \pmod{9}$
7	$8288256 \equiv 3 \pmod{9}$	$12 \equiv 3 \pmod{9}$	$-24 \equiv 3 \pmod{9}$
19	$392037661056 \equiv 3 \pmod{9}$	$12 \equiv 3 \pmod{9}$	$-24 \equiv 3 \pmod{9}$
31	$779869748441088 \equiv 0 \pmod{9}$	$36 \equiv 0 \pmod{9}$	$-72 \equiv 0 \pmod{9}$

Table 7.2. $p = 3$

D	$\dim W_D$	$\mathrm{tr}(g_5\|W_D)$	$-24H(D)$
3	$26752 \equiv 2 \pmod{5}$	$2 \equiv 2 \pmod{5}$	$-8 \equiv 2 \pmod{5}$
7	$8288256 \equiv 1 \pmod{5}$	$6 \equiv 1 \pmod{5}$	$-24 \equiv 1 \pmod{5}$
23	$6103910176768 \equiv 3 \pmod{5}$	$18 \equiv 3 \pmod{5}$	$-72 \equiv 3 \pmod{5}$
47	$2548919136928993280 \equiv 0 \pmod{5}$	$30 \equiv 0 \pmod{5}$	$-120 \equiv 0 \pmod{5}$

Table 7.3. $p = 5$

D	$\dim W_D$	$\mathrm{tr}(g_7\|W_D)$	$-24H(D)$
4	$143376 \equiv 2 \pmod{7}$	$2 \equiv 2 \pmod{7}$	$-12 \equiv 2 \pmod{7}$
8	$26124256 \equiv 4 \pmod{7}$	$4 \equiv 4 \pmod{7}$	$-24 \equiv 4 \pmod{7}$
11	$561346944 \equiv 4 \pmod{7}$	$4 \equiv 4 \pmod{7}$	$-24 \equiv 4 \pmod{7}$
15	$18508941312 \equiv 1 \pmod{7}$	$8 \equiv 1 \pmod{7}$	$-48 \equiv 1 \pmod{7}$
71	$4918685030138843868979 2 \equiv 0 \pmod{7}$	$28 \equiv 0 \pmod{7}$	$-168 \equiv 0 \pmod{7}$

Table 7.4. $p = 7$

7.2. **Selmer and Tate–Shafarevich Group Congruences.** Theorems 1.3 and 1.4 offer criteria for detecting elements in p-Selmer groups and Tate–Shafarevich groups of quadratic twists of certain elliptic curves. Theorem 1.3 assumes the truth of the Birch and Swinnerton-Dyer Conjecture. Theorem 1.4 is unconditional thanks to results of Skinner–Urban.

Here we offer data related to the curves E_{14} and E_{15}. In the notation of Theorem 1.4, we consider fundamental discriminants $-D$ such that $\left(\frac{-D}{p}\right) = -1$ and $\left(\frac{-D}{p'}\right) = 1$. For convenience let

$$\begin{aligned} H_{14}(D) &:= \delta_7(H(D) - \delta_7 H^{(2)}(D)), \\ \mathrm{tr}_2(D) &:= \mathrm{tr}(g_2|W_D), \\ \mathrm{Diff}_{14}(D) &:= H_{14}(D) - \mathrm{tr}_2(D), \end{aligned} \qquad \begin{aligned} H_{15}(D) &:= \delta_5(H(D) - \delta_5 H^{(3)}(D)), \\ \mathrm{tr}_3(D) &:= \mathrm{tr}(g_3|W_D), \\ \mathrm{Diff}_{15}(D) &:= H_{15}(D) - \mathrm{tr}_3(D). \end{aligned}$$

We have the following numerics. In Tables 7.5 and 7.6, the second and third columns offer graded traces and differences of class numbers. The fourth and fifth columns offer Mordell–Weil ranks and orders of Tate–Shafarevich groups assuming the Birch and Swinnerton-Dyer Conjecture. By Theorem 1.4, these columns are congruent if and only the corresponding p-Selmer group is nontrivial. First note that if these two columns are incongruent, then both the Mordell–Weil rank over $\mathbb{Q}$ and the p-part of the Tate–Shafarevich groups are trivial. However, when these columns are congruent, notice that either the rank is positive or the Tate–Shafarevich group is nontrivial at p.

D	$\mathrm{tr}_2(D)$	$H_{14}(D)$	$\mathrm{Diff}_{14}(D) \pmod 7$	$\mathrm{rk}(E_{14}(-D))$	$\#Ш_{an}(E_{14}(-D))$
15	-96256	-30	3	0	1
23	-1746944	-45	0	2	1
39	-165767168	-60	4	0	1
71	-156822906880	-105	4	0	1
79	-669595144192	-75	3	0	1
239	-6190369...040	-225	0	2	1
2671	-1630362...664	-345	0	0	49

TABLE 7.5. Examples for the curve E_{14}

D	$\mathrm{tr}_3(D)$	$H_{15}(D)$	$\mathrm{Diff}_{15}(D) \pmod 5$	$\mathrm{rk}(E_{15}(-D))$	$\#Ш_{an}(E_{15}(-D))$
8	-188	-6	3	0	1
23	-11456	-18	2	0	1
47	-860032	-30	3	0	1
68	-15834144	-24	0	2	1
83	-96763256	-18	2	0	1
248	-10546706...288	-48	0	2	1
308	-45931281...288	-48	0	2	1
587	-54506997...592	-42	0	0	25
1523	-15706167...792	-42	0	0	25

Table 7.6. Examples for the curve E_{15}

The authors thank Drew Sutherland for computing the elliptic curve invariants in Tables 7.5 and 7.6.

Appendix A. The Character Table of *O'N*

Here we give the character table of the O'Nan group $O'N$ over the complex numbers. For $n \in \mathbb{N}$ we let $\zeta_n := e^{\frac{2\pi i}{n}}$ and define

$$A := \frac{1+3\sqrt{5}}{2}, \qquad B := \sqrt{2},$$
$$C := -\zeta_{19} - \zeta_{19}^{7} - \zeta_{19}^{8} - \zeta_{19}^{11} - \zeta_{19}^{12} - \zeta_{19}^{18},$$
$$D := -\zeta_{19}^{4} - \zeta_{19}^{6} - \zeta_{19}^{9} - \zeta_{19}^{10} - \zeta_{19}^{13} - \zeta_{19}^{15},$$
$$E := -\zeta_{19}^{2} - \zeta_{19}^{3} - \zeta_{19}^{5} - \zeta_{19}^{14} - \zeta_{19}^{16} - \zeta_{19}^{17},$$
$$F := i\sqrt{5}, \qquad G := \sqrt{7}, \qquad H := \frac{-1+i\sqrt{31}}{2}.$$

We use $\overline{A}$, $\overline{B}$, &c. to denote images under the obvious Galois involutions. Note that C, D, and E are in one Galois orbit as well, since

$$(X-C)(X-D)(X-E) = X^3 - X^2 - 6X + 7.$$

The character table is reproduced from Gap4 [50].

	$1A$	$2A$	$3A$	$4A$	$4B$	$5A$	$6A$	$7A$	$7B$	$8A$	$8B$	$10A$	$11A$	$12A$	$14A$	$15A$	$15B$	$16A$	$16B$	$16C$	$16D$	$19A$	$19B$	$19C$	$20A$	$20B$	$28A$	$28B$	$31A$	$31B$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	10944	64	9	64	0	-1	1	17	3	0	0	-1	-1	1	1	-1	-1	0	0	0	0	0	0	0	-1	-1	1	1	1	1
χ_3	13376	-64	11	64	0	1	-1	-1	-1	0	0	1	0	1	-1	1	1	0	0	0	0	0	0	0	-1	-1	1	1	H	$\overline{H}$
χ_4	13376	-64	11	64	0	1	-1	-1	-1	0	0	1	0	1	-1	1	1	0	0	0	0	0	0	0	-1	-1	1	1	$\overline{H}$	H
χ_5	25916	-36	-4	20	4	1	0	-5	2	0	0	-1	0	2	-1	1	1	0	0	0	0	0	0	0	F	$-F$	-1	-1	0	0
χ_6	25916	-36	-4	20	4	1	0	-5	2	0	0	-1	0	2	-1	1	1	0	0	0	0	0	0	0	$-F$	F	-1	-1	0	0
χ_7	26752	128	22	0	0	2	2	-2	-2	0	0	-2	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	-1	-1
χ_8	32395	75	-5	35	3	0	3	6	-1	3	-1	0	0	-1	-2	0	0	1	1	-1	-1	0	0	0	0	0	0	0	0	0
χ_9	32395	75	-5	35	3	0	3	6	-1	-1	3	0	0	-1	-2	0	0	-1	-1	1	1	0	0	0	0	0	0	0	0	0
χ_{10}	37696	-64	31	-64	0	1	-1	15	1	0	0	1	-1	-1	-1	1	1	0	0	0	0	0	0	0	1	1	-1	-1	0	0
χ_{11}	52668	92	18	20	4	-2	2	-7	0	0	0	2	0	2	1	-2	-2	0	0	0	0	0	0	0	0	0	-1	-1	-1	-1
χ_{12}	58311	71	-9	71	7	1	-1	1	1	-1	-1	1	0	-1	1	1	1	-1	-1	-1	-1	0	0	0	1	1	1	1	0	0
χ_{13}	58311	71	-9	-1	-1	1	-1	1	1	3	-1	1	0	-1	1	1	1	-1	-1	1	1	0	0	0	-1	-1	-1	-1	0	0
χ_{14}	58311	71	-9	-1	-1	1	-1	1	1	-1	3	1	0	-1	1	1	1	1	1	-1	-1	0	0	0	-1	-1	-1	-1	0	0
χ_{15}	58653	-35	9	-35	-3	8	1	0	0	1	1	0	1	1	0	-1	-1	-1	-1	-1	-1	0	0	0	0	0	0	0	1	1
χ_{16}	64790	70	-10	-70	2	0	-2	12	-2	0	0	0	0	2	0	0	0	B	$-B$	$-B$	B	0	0	0	0	0	0	0	0	0
χ_{17}	64790	70	-10	-70	2	0	-2	12	-2	0	0	0	0	2	0	0	0	$-B$	B	B	$-B$	0	0	0	0	0	0	0	0	0
χ_{18}	85064	-56	14	-56	8	-1	-2	0	0	0	0	-1	1	-2	0	-1	-1	0	0	0	0	1	1	1	-1	-1	0	0	0	0
χ_{19}	116963	35	-1	35	3	8	-1	0	0	-1	-1	0	0	-1	0	-1	-1	1	1	1	1	-1	-1	-1	0	0	0	0	0	0
χ_{20}	143374	14	4	126	-2	-1	-4	0	0	2	2	-1	0	0	0	-1	-1	0	0	0	0	0	0	0	1	1	0	0	-1	-1
χ_{21}	169290	90	0	-90	-2	0	0	-5	2	0	0	0	0	0	-1	0	0	B	$-B$	B	$-B$	0	0	0	0	0	1	1	-1	-1
χ_{22}	169290	90	0	-90	-2	0	0	5	2	0	0	0	0	0	-1	0	0	$-B$	B	$-B$	B	0	0	0	0	0	1	1	-1	-1
χ_{23}	175616	0	8	0	0	-4	0	0	0	0	0	0	1	0	0	A	$\overline{A}$	0	0	0	0	-1	-1	-1	0	0	0	0	1	1
χ_{24}	175616	0	8	0	0	-4	0	0	0	0	0	0	1	0	0	$\overline{A}$	A	0	0	0	0	-1	-1	-1	0	0	0	0	1	1
χ_{25}	175770	90	0	90	-6	0	0	7	0	2	-2	0	1	0	-1	0	0	0	0	0	0	1	1	1	0	0	-1	-1	0	0
χ_{26}	207360	0	0	0	0	0	0	-8	-1	0	0	0	-1	0	0	0	0	0	0	0	0	C	E	D	0	0	0	0	1	1
χ_{27}	207360	0	0	0	0	0	0	-8	-1	0	0	0	-1	0	0	0	0	0	0	0	0	D	C	E	0	0	0	0	1	1
χ_{28}	207360	0	0	0	0	0	0	-8	-1	0	0	0	-1	0	0	0	0	0	0	0	0	E	D	C	0	0	0	0	1	1
χ_{29}	234080	-160	-10	0	0	0	2	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	G	$-G$	-1	-1
χ_{30}	234080	-160	-10	0	0	0	2	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	$-G$	G	-1	-1

TABLE A.1. Character table of *O’N*

APPENDIX B. MULTIPLICITIES OF IRREDUCIBLE REPRESENTATIONS IN W

We denote by V_j the $O'N$-module corresponding to the irreducible χ_j in Table A.1.

The following table gives the multiplicities of V_j in the (virtual) modules W_m in Theorem 4.1. Negative multiplicities are printed in bold.

m	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}
3	0	0	0	0	0	0	1	0	0	0
4	1	0	0	0	0	0	0	0	0	0
7	0	0	1	1	1	1	**-2**	0	0	2
8	**-2**	1	0	0	2	2	0	2	2	1
11	0	18	8	8	28	28	40	48	48	34
12	**-1**	33	44	44	76	76	88	98	98	122
15	0	406	581	581	1061	1061	1010	1252	1252	1568
16	**-2**	978	1193	1193	2316	2316	2386	2892	2892	3362
19	2	9484	11205	11205	21948	21948	23114	27766	27766	31894
20	5	18951	23161	23161	44930	44930	46322	56156	56156	65271
23	2	144238	177831	177831	343685	343685	352892	428308	428308	499900
24	25	277191	338794	338794	656282	656282	677588	820362	820362	954783
27	212	1795740	2189365	2189365	4245047	4245047	4388491	5310882	5310882	6174470
28	292	3264537	3989983	3989983	7730566	7730566	7979966	9663217	9663217	11244510
31	1562	18513448	22644956	22644956	43863830	43863830	45258570	54815104	54815104	63803360
32	2960	32416998	39620773	39620773	76765848	76765848	79241546	95957290	95957290	111658534
35	15432	165271652	201946677	201946677	391304807	391304807	403986962	489174874	489174874	569165006
36	25645	279985728	342204752	342204752	663020690	663020690	684409504	828775828	828775828	964395212

TABLE B.1. Multiplicities, part I.

m	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}	V_{16}	V_{17}	V_{18}	V_{19}	V_{20}
3	0	0	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	1	0	0
7	0	0	0	0	2	0	0	2	2	2
8	2	2	4	4	3	4	4	2	7	8
11	72	80	80	80	64	88	88	96	144	176
12	164	173	178	178	185	197	197	261	359	444
15	2068	2296	2296	2296	2384	2556	2556	3458	4680	5754
16	4704	5210	5200	5200	5224	5782	5782	7598	10432	12788
19	45058	49802	49802	49802	49804	55314	55314	72214	99604	122014
20	91248	101087	101068	101068	101628	112302	112302	147407	202710	248454
23	696576	771644	771644	771644	777260	857476	857476	1127304	1548902	1898946
24	1333868	1476646	1476680	1476680	1485435	1640744	1640744	2154259	2962056	3630946
27	8633536	9557140	9557140	9557140	9609292	10618702	10618702	13936084	19166220	23493012
28	15710534	17393783	17393848	17393848	17495880	19326474	19326474	25374046	34889380	42767664
31	89122420	98675012	98675012	98675012	99266748	109640000	109640000	143966514	197940198	242639964
32	156007392	172723134	172723024	172723024	173735642	191914504	191914504	251967626	346455808	424687592
35	795291752	880491400	880491400	880491400	885616006	978320518	978320518	1284400160	1766091974	2164876128
36	1347430236	1491796517	1491796318	1491796318	1500546542	1657551544	1657551544	2176231689	2992317414	3668002182

TABLE B.2. Multiplicities, part II.

m	V_{21}	V_{22}	V_{23}	V_{24}	V_{25}	V_{26}	V_{27}	V_{28}	V_{29}	V_{30}
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
7	2	2	3	3	2	4	4	4	6	6
8	10	10	9	9	10	12	12	12	14	14
11	216	216	214	214	224	252	252	252	270	270
12	521	521	543	543	542	638	638	638	718	718
15	6746	6746	7057	7057	7006	8328	8328	8328	9492	9492
16	15102	15102	15671	15671	15680	18500	18500	18500	20884	20884
19	144268	144268	149402	149402	149782	176414	176414	176414	198703	198703
20	293374	293374	304323	304323	304596	359352	359352	359352	405676	405676
23	2241422	2241422	2326161	2326161	2327256	2746666	2746666	2746666	3102368	3102368
24	4287248	4287248	4447476	4447476	4451367	5251357	5251357	5251357	5927992	5927992
27	27742332	27742332	28775511	28775511	28804106	33976834	33976834	33976834	38348849	38348849
28	50498270	50498270	52385258	52385258	52431245	61854317	61854317	61854317	69824744	69824744
31	286490080	286490080	297207048	297207048	297456630	350929578	350929578	350929578	396169260	396169260
32	501453364	501453364	520191449	520191449	520647692	614220424	614220424	614220424	693367868	693367868
35	2556221884	2556221884	2651707883	2651707883	2654066434	3131025718	3131025718	3131025718	3534425359	3534425359
36	4331022760	4331022760	4492864127	4492864127	4496803456	5304984880	5304984880	5304984880	5988574304	5988574304

Table B.3. Multiplicities, part III.

Appendix C. Congruences

$\underline{p = 31}$:

$$0 \equiv F_{1A} - F_{31AB} \pmod{31}$$

$\underline{p = 19}$:

$$0 \equiv F_{1A} - F_{19ABC} \pmod{19}$$

$\underline{p = 11}$:

$$0 \equiv F_{1A} - F_{11A} \pmod{11}$$

$\underline{p = 7}$:

$$\begin{aligned} 0 &\equiv F_{1A} - F_{7AB} && \pmod{7^3} \\ &\equiv F_{2A} - F_{14A} && \pmod{7} \\ &\equiv F_{4AB} - F_{28AB} && \pmod{7} \end{aligned}$$

$\underline{p = 5}$:

$$\begin{aligned} 0 &\equiv F_{1A} - F_{5A} && \pmod{5^3} \\ &\equiv F_{2A} - F_{10A} && \pmod{5} \\ &\equiv F_{3A} - F_{15AB} && \pmod{5} \\ &\equiv F_{4AB} - F_{20AB} && \pmod{5} \end{aligned}$$

$\underline{p=3}$:

$$\begin{aligned}
0 &\equiv F_{1A} - F_{3A} && \pmod{3^5}\\
&\equiv F_{2A} - F_{6A} && \pmod{3^2}\\
&\equiv F_{4AB} - F_{12A} && \pmod{3^2}\\
&\equiv F_{5A} - F_{15AB} && \pmod{3^2}
\end{aligned}$$

$\underline{p=2}$:

$$\begin{aligned}
0 &\equiv F_{1A} + 303F_{2A} + 3024F_{4AB} + 4864F_{8AB} + 57344F_{16ABCD} && \pmod{2^{16}}\\
&\equiv F_{2A} + 7F_{4AB} + 8F_{8AB} + 112F_{16ABCD} && \pmod{2^7}\\
&\equiv F_{3A} + F_{6A} + 6F_{12A} && \pmod{2^3}\\
&\equiv F_{4AB} + F_{8AB} + 14F_{16ABCD} && \pmod{2^4}\\
&\equiv F_{5A} + F_{10A} + 6F_{20AB} && \pmod{2^3}\\
&\equiv F_{6A} + F_{12A} && \pmod{2}\\
&\equiv F_{7AB} + F_{14AB} && \pmod{2^3}\\
&\equiv F_{8AB} + 7F_{16ABCD} && \pmod{2^3}\\
&\equiv F_{10A} + F_{20AB} && \pmod{2}\\
&\equiv F_{14A} + F_{28AB} && \pmod{2}
\end{aligned}$$

Appendix D. Traces of Singular Moduli

We give the explicit descriptions of $F_{[g]}$ in terms of traces of singular moduli and class numbers as described in Section 5.

$$\begin{aligned}
F_{1A} &= \mathscr{T}^{(1)},\\
F_{2A} &= \mathscr{T}^{(2)} + 12\mathscr{H}^{(1)} - 12\mathscr{H}^{(2)},\\
F_{3A} &= \mathscr{T}^{(3)} + 12\mathscr{H}^{(1)} - 12\mathscr{H}^{(3)},\\
F_{4AB} &= \mathscr{T}^{(4)} + 12\mathscr{H}^{(2)} - 12\mathscr{H}^{(4)},\\
F_{5A} &= \mathscr{T}^{(5)} + 6\mathscr{H}^{(1)} - 6\mathscr{H}^{(5)},\\
F_{6A} &= \mathscr{T}^{(6)} - 12\mathscr{H}^{(1)} + 8\mathscr{H}^{(2)} + \frac{21}{2}\mathscr{H}^{(3)} - \frac{13}{2}\mathscr{H}^{(6)},\\
F_{7AB} &= \mathscr{T}^{(7)} + 4\mathscr{H}^{(1)} - 4\mathscr{H}^{(7)},\\
F_{8AB} &= \mathscr{T}^{(8)} + 4\mathscr{H}^{(4)} - 4\mathscr{H}^{(8)},\\
F_{10A} &= \mathscr{T}^{(10)} - 6\mathscr{H}^{(1)} + 4\mathscr{H}^{(2)} + \frac{11}{2}\mathscr{H}^{(5)} - \frac{7}{2}\mathscr{H}^{(10)},\\
F_{11A} &= \mathscr{T}^{(11,+)} + \frac{12}{5}\mathscr{H}^{(1)} - \frac{6}{5}\mathscr{H}^{(11)} - \frac{4}{5}\mathscr{G}^{(11)},\\
F_{12A} &= \mathscr{T}^{(12)} - 4\mathscr{H}^{(2)} + 4\mathscr{H}^{(4)} + \frac{5}{2}\mathscr{H}^{(6)} - \frac{5}{2}\mathscr{H}^{(12)},\\
F_{14A} &= \mathscr{T}^{(14,+)} - 4\mathscr{H}^{(1)} + \frac{8}{3}\mathscr{H}^{(2)} + \frac{15}{4}\mathscr{H}^{(7)} - \frac{41}{24}\mathscr{H}^{(14)} + \frac{8}{3}\mathscr{G}^{(14)},\\
F_{15AB} &= \mathscr{T}^{(15,+)} - 3\mathscr{H}^{(1)} + \frac{9}{4}\mathscr{H}^{(3)} + \frac{5}{2}\mathscr{H}^{(5)} - \frac{13}{8}\mathscr{H}^{(15)} + \frac{9}{4}\mathscr{G}^{(15)},\\
F_{16ABCD} &= 2\mathscr{T}^{(32,+)} - \mathscr{T}^{(16)} - 2\mathscr{H}^{(8)} + 4\mathscr{H}^{(16)} - \mathscr{H}^{(32)},\\
F_{19ABC} &= \mathscr{T}^{(19,+)} + \frac{4}{3}\mathscr{H}^{(1)} - \frac{2}{3}\mathscr{H}^{(19)} + \frac{4}{3}\mathscr{G}^{(19)},\\
F_{20AB} &= \mathscr{T}^{(20,+)} - 2\mathscr{H}^{(2)} + 2\mathscr{H}^{(4)} + \frac{3}{2}\mathscr{H}^{(10)} - \frac{3}{2}\mathscr{H}^{(20)},\\
F_{28AB} &= \mathscr{T}^{(28,+)} - \frac{4}{3}\mathscr{H}^{(2)} + \frac{4}{3}\mathscr{H}^{(4)} + \frac{25}{24}\mathscr{H}^{(14)} - \frac{25}{24}\mathscr{H}^{(28)} + \frac{8}{3}\mathscr{G}^{(28)},\\
F_{31AB} &= \mathscr{T}^{(31,+)} + \frac{4}{5}\mathscr{H}^{(1)} - \frac{2}{5}\mathscr{H}^{(31)} + \frac{3}{5}\mathscr{G}^{(31)}.
\end{aligned}$$

Here, $\mathscr{G}^{(N)}$ denotes the unique weight 3/2 cusp form for $\Gamma_0(4N)$ in the plus space with leading coefficient 1 if $N < 28$, $\mathscr{G}^{(28)}$ is the unique normalized cusp form in $S_{\frac{3}{2}}(\Gamma_0(28))$ hit with the V_4-operator, and $\mathscr{G}^{(31)} \in S_{\frac{3}{2}}^{+}(\Gamma_0(124))$ is the unique cusp form in this space satisfying

$$\mathscr{G}^{(31)}(\tau) = q^4 + \frac{11}{3}q^7 + O(q^8).$$

There is exactly one newform $f^{(31)}$ in $S^{+}_{\frac{3}{2}}(\Gamma_0(124))$ up to Galois conjugation and we have

$$f^{(31)}(\tau) = q^4 - \frac{1+\sqrt{5}}{2}q^7 - \frac{1-\sqrt{5}}{2}q^8 + \frac{1+\sqrt{5}}{2}q^{16} + O(q^{20}),$$

so that we can express $g^{(31)}$ in terms of this newform as follows,

$$\mathscr{G}^{(31)}(\tau) = \frac{3+5\sqrt{5}}{6}f^{(31)}(\tau) + \frac{3-5\sqrt{5}}{6}f^{(31)^{\sigma}}(\tau),$$

where, as in Table 5.2, a superscript σ denotes Galois conjugation.

References

[1] A. Adem, J. Carlson, D. Karagueuzian and R. Milgram, *The cohomology of the Sylow 2-subgroup of the Higman-Sims group*, J. Pure Appl. Algebra **164** (2001), no. 3, 275–305.

[2] A. Adem and R. Milgram, *The subgroup structure and mod 2 cohomology of O'Nan's sporadic simple group*, J. Algebra **176** (1995), no. 1, 288–315.

[3] A. Agashe, *A visible factor of the special L-value*, J. Reine Angew. Math. **644** (2010), 159–187.

[4] ———, *Squareness of the special L-values of twists*, Int. J. Number Theory **6** (2010), no. 5, 1091–1111.

[5] J. Alperin, *Sylow 2-subgroups of 2-rank three*, Finite groups '72 (Proc. Gainesville Conf., Univ. Florida, Gainsville, Fla., 1972), 3–5. North-Holland Math. Studies, Vol. 7, North-Holland, Amsterdam, 1973.

[6] V. Anagiannis, M. Cheng and S. Harrison, *K3 Elliptic Genus and an Umbral Moonshine Module*, preprint, available at `https://arxiv.org/abs/1709.01952`, 2017.

[7] M. Aschbacher, *The status of the classification of the finite simple groups*, Notices Amer. Math. Soc. **51** (2004), no. 7, 736–740. MR 2072045

[8] L. Beneish and H. Larson, *Traces of singular values of hauptmoduln*, Int. J. Number Theory **11** (2015), 1027–1048.

[9] M. Bhargava and A. Shankar, *Ternary cubic forms having bounded invariants and the proof of a positive proportion of elliptic curves having rank* 0, Ann. Math. **181** (2015), no. 2, 587–621.

[10] R. Borcherds, *Monstrous moonshine and monstrous Lie superalgebras*, Invent. Math. **109** (1992), no. 2, 405–444.

[11] ———, *Automorphic forms on $O_{s+2,2}(\mathbb{R})$ and infinite products*, Invent. Math. **120**, (1995), 161-213.

[12] ———, *Problems in moonshine*, First International Congress of Chinese Mathematicians, (Beijing, 1998), AMS/IP Stud. Adv. Math., 20, Amer. Math. Soc., Providence, RI, 2001.

[13] W. Bosma, J. Cannon, and C. Playoust. *The Magma algebra system. I The user language*, J. Symbolic Comput., **24** (1997), no. 3–4, 235–265.

[14] K. Bringmann, A. Folsom, K. Ono, and L. Rolen, *Harmonic Maass Forms and Mock Modular Forms: Theory and Applications*, Colloquium publications, Amer. Math. Soc. (2017), to appear.

[15] K. Bringmann and K. Ono, *The $f(q)$ mock theta function conjecture and partition ranks*, Invent. Math. **165** (2006), 243–266.

[16] ———, *Arithmetic properties of coefficients of half-integral weight harmonic Maass forms*, Math. Ann. **337** (2007), 591–612.

[17] ———, *Dyson's ranks and Maass forms*, Ann. Math. **171** (2010), 419–449.

[18] J. Bruinier and J. Funke, *On two geometric theta lifts*, Duke Math. J. **1** (2004), no. 125, 45–90.

[19] ———, *Traces of CM values of modular functions*, J. Reine Angew. Math. **594** (2006), 1–33.

[20] J. Bruinier and K. Ono, *Heegner divisors, L-functions, and harmonic weak Maass forms*, Ann. Math. **172** (2010), 2135-2181.

[21] S. Carnahan, *Generalized Moonshine IV: Monstrous Lie Algebras*, preprint, available at `https://arxiv.org/abs/1208.6254`, 2012.

[22] M. Cheng and J. Duncan, *On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine*, Commun. Number Theory Physics **6** (3) (2012), 697–758.

[23] ———, *Rademacher Sums and Rademacher Series*, in W. Kohnen (ed.) and R. Weissauer (ed.) *Conformal Field Theory, Automorphic Forms and Related Topics: CFT, Heidelberg, September 19-23, 2011*, Springer-Verlag, 2014, 143–182.

[24] ———, *Optimal Mock Jacobi Theta Functions*, preprint, available at `https://arxiv.org/abs/1605.04480`, 2016.

[25] ———, *Meromorphic Jacobi Forms of Half-Integral Index and Umbral Moonshine Modules*, preprint, available at `https://arxiv.org/abs/1707.01336`, 2017.

[26] M. Cheng, J. Duncan, and J. Harvey, *Umbral Moonshine*, Commun. Number Theory Phys. **8** (2014), no. 2, 101–242.

[27] ———, *Umbral Moonshine and the Niemeier Lattices*, Res. Math. Sci. **1** (2014), Art. 3, 81 pp.

[28] ———, *Weight One Jacobi Forms and Umbral Moonshine*, J. Phys. A **51** (2018), no. 10, 104002, 37 pp.

[29] J. Conway and S. Norton, *Monstrous Moonshine*, Bull. London Math. Soc. **11** (1979), 308–339.

[30] J. Cremonia, *Computing the degree of the modular parameterization of a modular elliptic curve*, Math. Comp. **64** (1995), 1235-1250.

[31] A. Dabholkar, S. Murthy, D. Zagier, *Quantum Black Holes, Wall Crossing, and Mock Modular Forms*, to appear in Cambridge Monographs in Mathematical Physics, available at `http://arxiv.org/abs/1208.4074`, 2012.

[32] W. Duke, *Hyperbolic distribution problems and half-integral weight Maass forms*, Invent. Math. **92** (1988), 73–90.

[33] W. Duke and R. Schulze-Pillot, *Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids*, Invent. Math. **99** (1990), 49–57.

[34] J. Duncan, *Moonshine for Rudvalis's sporadic group I*, arXiv:math/0609449.

[35] ———, *Moonshine for Rudvalis's sporadic group II*, arXiv:math/0611355.

[36] J. Duncan and I. Frenkel, *Rademacher sums, moonshine, and gravity*, Commun. Number Theory Phys. **5** (2011), no. 4, 1–128.

[37] J. Duncan, M. Griffin, and K. Ono, *Proof of the Umbral Moonshine Conjecture*, Res. Math. Sci. **2** (2015), Art. 26, 47 pp.

[38] ———, *Moonshine*, Res. Math. Sci. **2** (2015), Art. 11, 57 pp.

[39] J. Duncan and J. Harvey, *The Umbral Moonshine Module for the Unique Unimodular Niemeier Root System*, Algebra & Number Theory **11** (2017) no. 3, 505–535.

[40] J. Duncan and A. O'Desky, *Super vertex algebras, meromorphic Jacobi forms and umbral moonshine*, J. Algebra **515** (2018), 389–407.

[41] T. Eguchi, H. Ooguri, and Y. Tachikawa, *Notes on the K3 Surface and the Mathieu Group M_{24}*, Experiment. Math. **20** (2011), no. 1, 91–96.

[42] D. Evans and T. Gannon, *Reconstruction and Local Extensions for Twisted Group Doubles, and Permutation Orbifolds*, preprint, available at `https://arxiv.org/abs/1804.11145`, 2018.

[43] D. Ford, J. McKay, and S. Norton, *More on replicable functions*, Commun. Algebra **22** (1994), no. 13, 5175–5193.

[44] I. Frenkel, J. Lepowsky, and A. Meurman, *A natural representation of the Fischer-Griess Monster with the modular function J as character*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 10, Phys. Sci., 3256–3260.

[45] ———, *A moonshine module for the Monster*, Vertex operators in mathematics and physics (Berkeley, Calif., 1983), Math. Sci. Res. Inst. Publ., vol. 3, Springer, New York, 1985, 231–273.

[46] ———, *Vertex operator algebras and the Monster*, Pure and Applied Mathematics, vol. 134, Academic Press Inc., Boston, MA, 1988.

[47] T. Gannon, *Monstrous moonshine: the first twenty-five years*, Bull. London Math. Soc. **38** (2006), no. 1, 1–33.

[48] ———, *Moonshine beyond the Monster*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2006.

[49] ———, *Much ado about Mathieu*, Adv. Math. **301** (2016), 322–358.

[50] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.7.8;* 2016, `http://www.gap-system.org`.

[51] D. Goldfeld, *Conjectures on elliptic curves over quadratic fields*, in Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), M. B. Nathanson, ed., Lect. Notes in Math. **751**, Springer, Berlin, 1979, 108–118.

[52] R. Griess Jr., *The friendly giant*, Invent. Math. **69** (1982), no. 1, 1–102.

[53] ———, *Sporadic groups, code loops and nonvanishing cohomology*, Proceedings of the Northwestern conference on cohomology of groups (Evanston, Ill., 1985). J. Pure Appl. Algebra **44** (1987), no. 1-3, 191–214.

[54] M. Griffin and M. Mertens, *A proof of the Thompson Moonshine Conjecture*, Res. Math. Sci. **3** (2016), Art. 36, 32 pp.

[55] B. Gross, *The classes of singular moduli in the generalized Jacobian*, Geometry and arithmetic (Ed. C. Faber et. al.), European Math. Soc., Zürich (2012), 137–141.

[56] B. Gross, W. Kohnen, and D. Zagier, *Heegner Points and Derivatives of L-series. II*, Math. Ann. **278** (1987), 497–562.

[57] B. Gross and D. Zagier, *Heegner points and derivatives of L-series*, Invent. Math. **85** (1986), 225–320.

[58] J. Harvey and B. Rayhaun, *Traces of Singular Moduli and Moonshine for the Thompson Group*, Commun. Number Theory Phys. **10** (2016), no. 1, 23-62.

[59] C. Hermite, *Sur la théorie des équations modulaires*, in Œuvres de Charles Hermite, volume 2 (French), edited and with a foreword by Émile Picard, Reprint of the 1908 original, Cambridge University Press (2009), 38–82.

[60] F. Hirzebruch and D. Zagier, *Intersection Numbers of Curves on Hilbert Modular Surfaces and Modular Forms of Nebentypus*, Invent. Math. **36** (1976), 57–113.

[61] L.-K. Hua, *Introduction to Number Theory*, Springer-Verlag Berlin Heidelberg New York, 1982.

[62] H. Iwaniec, *Fourier coefficients of modular forms of half-integral weight*, Invent. Math. **87** (1987), 385–401.

[63] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Colloquium Publications **53**, Amer. Math. Soc., Providence, RI., 2004.

[64] T. Johnson-Freyd, *personal communication.*

[65] T. Johnson-Freyd and D. Treumann, *Third homology of some sporadic finite groups*, preprint, available at `https://arxiv.org/abs/1810.00463`, 2018.

[66] W. Kohnen, *Fourier Coefficients of Modular Forms of Half-Integral Weight*, Math. Ann. **271** (1985), 237–268.

[67] V. Kolyvagin, *Finiteness of $E(\mathbb{Q})$ and $X(E, \mathbb{Q})$ for a class of Weil curves*, Math. USSR Izv. **32** (1989), 523–541.

[68] The LMFDB Collaboration, *The L-functions and Modular Forms Database*, `http://www.lmfdb.org`, 2016, [Online, accessed 31 Jan. 2017].

[69] M. Manickam, B. Ramakrishnan, and T. Vasudevan, *On Shintani correspondence*, Proc. Indian Acad. Sci. (Math Sci.) **99** (1989), 235–247.

[70] A. Miller and A. Pixton, *Arithmetic traces of non-holomorphic modular invariants*, Int. J. Number Theory **6** (2010), 69–87.

[71] S. Norton, *Generalized Moonshine*, Proc. Symp. Pure Math **47** (1987), 208–209.

[72] A. Ogg, *Automorphismes de courbes modulaires*, Séminaire Delange-PisotPoitou (16e année: 1974/75), Théorie des nombres, Fasc. 1, Exp. No. 7, Secrétariat Mathématique, Paris, 1975, p. 8.

[73] M. O'Nan, *Some evidence for the existence of a new simple group*, Proc. London Math. Soc. **32** (1976), 421-479.

[74] K. Ono, *Unearthing the visions of a master: harmonic Maass forms and number theory*, Current Developments in Mathematics **2008** (2009), 347–454.

[75] N. Paquette, D. Persson and R. Volpato, *Monstrous BPS-algebras and the superstring origin of moonshine*, Commun. Number Theory Phys. **10** (2016) no. 3, 433–526.

[76] ———, *BPS algebras, genus zero and the heterotic Monster*, J. Phys. A **50** (2017) no. 41.

[77] The PARI Group, PARI/GP version `2.11.0`, Univ. Bordeaux, 2018, `http://pari.math.u-bordeaux.fr/`.
[78] L. Queen, *Modular functions arising from some finite groups*, Math. Comp. **37** (1981), no. 156, 547–580.
[79] H. Rademacher, *The Fourier coefficients of the modular invariant* $J(\tau)$, Amer. J. Math. **60** (1938), 501–512.
[80] J. Rouse and J. Webb, *Spaces of modular forms spanned by eta-quotients*, Adv. Math. **272** (2015), 200–224.
[81] A. Ryba, *A new construction of the O'Nan simple group*, J. Algebra **112** (1988), 173–197.
[82] J.-P. Serre, *Proprietés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331.
[83] J.-P. Serre and H. Stark, *Modular Forms of Weight 1/2*, Modular Functions of One Variable VI (J.-P. Serre and D. B. Zagier, eds.), Lecture Notes in Mathematics, vol. 627, Springer Berlin Heidelberg, 1977, 27–67.
[84] J. Silverman, *The arithmetic of elliptic curves*, Springer, New York, 2009.
[85] C. Skinner, *Multiplicative reduction and the cyclotomic main conjecture for* GL_2, Pacific J. Math. **283** (2016), no. 1, 171–200.
[86] C. Skinner and E. Urban, *The Iwasawa main conjectures for* GL_2, Invent. Math. **195** (2014), no. 1, 1–277.
[87] J. Sturm, *On the congruence of modular forms*, Number Theory (New York, 1984-1985), Springer Lect. Notes. Math. **1240**, (1987), Springer, Berlin, 275-280.
[88] J. Tate, *The arithmetic of elliptic curves*, Invent. Math. **23** (1974), 179–206.
[89] J. Thompson, *Finite groups and modular functions*, Bull. London Math. Soc. **11** (1979), no. 3, 347–351.
[90] ______, *Some numerology between the Fischer-Griess Monster and the elliptic modular function*, Bull. London Math. Soc. **11** (1979), no. 3, 352–353.
[91] M. Ueda, *The decomposition of the spaces of cusp forms of half-integral weight and trace formula of Hecke operators*, J. Math. Kyoto Univ. **28** (1988), 505–555.
[92] M. Ueda and S. Yamana, *On newforms for Kohnen plus spaces*, Math. Z. **264** (2010), 1–13.
[93] J.-L. Waldspurger, *Correspondance de Shimura*, J. Math. Pures Appl. (9) **59** (1980), no. 1, 1–132.
[94] ______, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, J. Math. Pures Appl. (9) **60** (1981), no. 4, 375–484.
[95] R. Wilson, *Is* J_1 *a subgroup of the Monster?*, Bull. LMS 18 (1986), 349–350.
[96] D. Zagier, *Nombres de classes et formes modulaires de poids* $3/2$, C. R. Acad. Sci. Paris Sér. A-B **281** (1975), no. 21, Ai, A883–A886.
[97] ______, *Zetafunktionen und quadratische Körper - Eine Einführung in die höhere Zahlentheorie*, Springer-Verlag Berlin Heidelberg New York, 1981.
[98] ______, *Modular parameterizations of elliptic curves*, Canadian Math. Bull. **28** (1985), 372-384.
[99] ______, *Traces of singular moduli*, in "Motives, Polylogarithms and Hodge Theory" (Eds. F. Bogomolov, L. Katzarkov), Lecture Series **3**, International Press, Somerville (2002), 209–244.
[100] ______, *Ramanujan's mock theta functions and their applications [d'après Zwegers and Bringmann-Ono]*, Séminaire Bourbaki, 60ème année, 2006-2007, no. 986.
[101] S. Zwegers, *Mock Theta Functions*, Ph.D. thesis, Universiteit Utrecht, 2002.

Department of Mathematics and Computer Science, Emory University, 400 Dowman Drive, Atlanta, GA 30322
E-mail address: `john.duncan@emory.edu`

Mathematisches Institut der Universtät zu Köln, Weyertal 86-90, D-50931 Köln
E-mail address: `mmertens@math.uni-koeln.de`

Department of Mathematics and Computer Science, Emory University, 400 Dowman Drive, Atlanta, GA 30322
E-mail address: `ono@mathcs.emory.edu`

TASI Lectures on Moonshine

Vassilis Anagiannis*[1] and Miranda C. N. Cheng†[1,2]

[1]Institute of Physics, University of Amsterdam, Amsterdam, the Netherlands
[2]Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, the Netherlands

Abstract

Preliminary Version
Comments and suggestions are welcome.

The word moonshine refers to unexpected relations between the two distinct mathematical structures: finite group representations and modular objects. It is believed that the key to understanding moonshine is through physical theories with special symmetries. Recent years have seen a varieties of new ways in which finite group representations and modular objects can be connected to each other, and these developments have brought promises and also puzzles into the string theory community.

These lecture notes aim to bring graduate students in theoretical physics and mathematical physics to the forefront of this active research area. In Part II of this note, we review the various cases of moonshine connections, ranging from the classical monstrous moonshine established in the last century to the most recent findings. In Part III, we discuss the relation between the moonshine connections and physics, especially string theory. After briefly reviewing a recent physical realisation of monstrous moonshine, we will describe in some details the mystery of the physical aspects of umbral moonshine, and also mention some other setups where string theory black holes can be connected to moonshine.

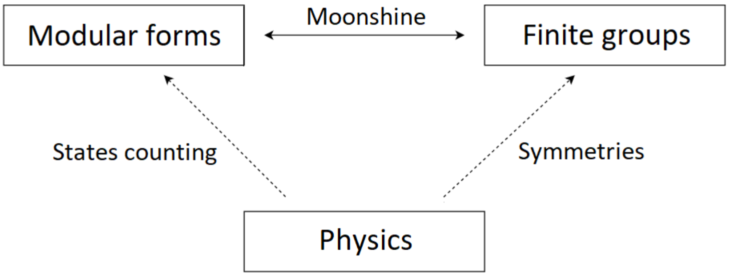
To make the exposition self-contained, we also provide the relevant background knowledge in Part I, including sections on finite groups, modular objects, and two-dimensional conformal field theories. This part occupies half of the pages of this set of notes and can be skipped by readers who are already familiar with the relevant concepts and techniques.

---

*v.anagiannis@uva.nl
†mcheng@uva.nl

Introduction

The word moonshine is employed in mathematics to refer to an unexpected relationship between modular objects and representations of finite groups. The study of moonshine phenomenon has seen rapid developments in the past five years. While the relation is between two mathematical structures, it is expected that the existence of this surprising relation has its origin in physics, and in string theory in particular.



Moonshine: what, how and why?

Q: What is moonshine?

The structures of modular forms and that of finite group representations have a priori nothing to do with each other. The word moonshine refers to an unexpected relation between them. But this simple answer calls for more questions. Where does moonshine occur? What types of modular objects feature in moonshine relations? What types of finite groups can be represented by (mock) modular forms? In what ways? Is there a classification possible of such moonshine relations? Before satisfying answers to these questions can be found, the existence of moonshine relation remains to a large extent a mysterious phenomenon.

Q: How do moonshine relations arise?

In the classical cases of monstrous (and similarly Conway) moonshine, reviewed in §4, the relevant group representation has the structure of vertex operator algebra (VOA). In physical terms, these cases of moonshine can be can be thought of as being "explained" by the existence of certain special 2-dimensional conformal field theories (CFTs) that have the relevant finite groups as symmetry groups.

It would be very gratifying to have a similar physical construction for the finite group modules underlying umbral and other recently discovered moonshine relations. In certain "simpler" cases this has been achieved (cf. §5.2). Nevertheless, a uniform construction of the umbral moonshine module, reflecting the uniform structure among the twenty-three instances of umbral moonshine, is still absent. Similarly, the modules underlying the various other cases of new moonshines have not been constructed so far.

Apart from the question "what are these finite group modules?" there is also a deeper question of "what kind of (algebraic) structure do these moonshine modules possess?". The *mock* modular nature of the modular objects involved in the new cases of moonshine seems to suggest that a departure from the familiar unitary conformal field theories with discrete spectrum is necessary in order to accommodate the corresponding moonshine modules. However, the existence of generalised umbral moonshine (cf. §5.2) suggests that certain features resembling those of holomorphic orbifold CFTs should still be present in these new modules.

Last but not least, the first case of umbral moonshine, often referred to as the Mathieu moonshine, was uncovered in the context of string theory in the background of $K3$ surfaces. Subsequent development established the intimate relation between $K3$ string theory and all twenty-three instances of umbral moonshine. This relation to $K3$ string theory is the topic of §8.

To sum up, the origin of the classical monstrous moonshine can be understood as lying in the existence of certain special physical theory – certain 2d chiral CFTs and the corresponding string theory embedding to be more precise. See §4 and §7. The search for the origin of the newer cases of moonshine is still largely a challenging puzzle. However, one does expect the answer to again lie in certain physical theories, and most probably arises from the framework of string theory.

Q: Why should I care about moonshine?

Here are a few reasons why the authors, and many other mathematicians and physicists, care about moonshine phenomenon. First, theoretical physics is to a large extent about symmetries, and moonshine is to a large extent about hidden symmetries, which (conjecturally) take place in a physical context. There are therefore very good reasons for physicists to care about what is going on in moonshine. More generally speaking, the broad and interdisciplinary nature of the question and its connection to different areas in mathematics and theoretical physics suggests that the understanding of this mathematical paradigm is likely to spur novel developments in these areas, as was clearly true in the case of the classical monstrous moonshine. Here we highlight a few examples of such connections, in the context of recent developments:

- The discovery of umbral moonshine was initiated in the context of $K3$ string theory, arguably one of the most important examples of string theory compactifications. The development of umbral moonshine has led to new results in the study of automorphic forms and $K3$ geometry, especially in the context of string theory. We expect it to also shed light on the structure of BPS states and on other interesting aspects of the string landscape in the future. See §5.1, 5.2 and §8.

- The discovery of various new moonshine examples (cf. §6), involving various types of finite groups including an infinite family of them, poses a fascinating challenge for theoretical physics to accommodate the corresponding group representations. The solution of this (completely well-posed) puzzle is likely to point to novel structures in physics.

- The connection between certain new examples of moonshine and arithmetic geometry is interesting and constitutes a potentially fruitful path towards new perspectives in number theory. See §6.

Moreover, it is fascinating and puzzling why the two distinct structures of modular objects and finite groups are so deeply connected. An understanding of the true nature of these connections has the potential to change the way mathematicians think about these objects in a profound way.

About these lecture notes

The lecture note is written for the occasion of the TASI 2017 lectures on Moonshine given by the second author.

Acknowledgement

The second author would like to thank the organisers of the TASI 2017 school for giving her the opportunity to lecture in this very nice school. We would like to thank OUR FRIEND**A**, **B**..... for comments on an earlier version of the draft. Last but not least, the second author would like to thank Sarah Harrison, Jeff Harvey, Shamit Kachru, Natalie Paquette, Brandon Rayhaun, Roberto Volpato, Max Zimet, and in particular John Duncan for enjoyable collaborations on related topics, and for having taught me so much about this really fun subject!

Contents

I Background **6**

1 Finite groups and their representations **6**
1.1 Groups 6
1.2 Classification of finite groups 8
1.3 Sporadic groups and lattices 9
1.4 Representations of finite groups 12

2 Modular objects **15**
2.1 Modular forms 15
2.2 (Skew-)Holomorphic Jacobi forms 19
2.3 Mock modular forms 22

3 Conformal field theory in 2 dimensions **25**
3.1 General structure 26
3.2 Orbifolds 30
3.3 Elliptic genus 33

II Moonshine **38**

4 Moonshine at weight zero **38**
4.1 Monstrous moonshine 38
4.2 Conway moonshine 43

5 Moonshine at weight one-half **46**
5.1 Mathieu moonshine 46
5.2 Umbral moonshine 48
5.3 Thompson moonshine 53

6 Moonshine at weight three-halves **55**

III Moonshine and string theory **58**

7 Monstrous moonshine **58**

8 Moonshine and $K3$ string theory **59**
8.1 Mathieu moonshine and $K3$ elliptic genus 60
8.2 Umbral moonshine and $K3$ elliptic genus 62
8.3 Conway moonshine and $K3$ elliptic genus 63
8.4 Symmetries of $K3$ string theory and moonshine 64
8.5 Beyond perturbative string theory 66

9 Other connections **68**

References **69**

Part I
Background

In the first part we briefly summarise the relevant background knowledge in moonshine study, including sections on finite groups, modular objects, and two-dimensional conformal field theories. This part can be safely skipped by readers who are familiar with these topics.

1 Finite groups and their representations

In this chapter we briefly review of basic notions of finite groups and their representations. We mainly follow [1–3] (see also [4]).

1.1 Groups

A group is a set G, together with a "multiplication" operation $\bullet : G \times G \to G$, formally denoted as $(G, \bullet)$. The symbol for this operation is usually implicit, and we often write ab for $a \bullet b$. A group must satisfy the following axioms:

1. **Closure:** $ab = c \in G$ for any $a, b \in G$.
2. **Associativity:** $(ab)c = a(bc)$ for any $a, b, c \in G$.
3. **Identity:** There exists a unique identity element $e \in G$, such that $eg = ge = g$ for any $g \in G$.
4. **Inverses:** For every $g \in G$, there exists a unique inverse element $g^{-1} \in G$, such that $gg^{-1} = g^{-1}g = e$. We also have that $e^{-1} = e$.

Notice that $ab \neq ba$ in general. In the case that $ab = ba$ for every $a, b \in G$, i.e. the group operation is commutative, the group is called *Abelian.* The number of elements of G is called the *order* of G, and it can be either finite or infinite. We also define the *order of an element*, $|g|$, to be the minimum number of times we need to multiply it with itself in order to reach the identity, i.e. $g^{|g|} = e$ (the order can also be infinite).

Next we give a brief summary of a few important notions of group theory.

Group homomorphisms. We say that a map $\phi : G \to F$ between two groups $(G, \bullet)$ and $(F, \star)$ is a *group homomorphism* if it preserves the group structure of G. In other words, ϕ must satisfy

$$\phi(a \bullet b) = \phi(a) \star \phi(b) \ , \tag{1.1}$$

for all $a, b \in G$. If there also exists the inverse homomorphic map $\phi^{-1} : F \to G$, then G and F are isomorphic; such groups are abstractly the same, but they may still have different realisations. An isomorphism $G \to G$ is called *automorphism*, and is often called a symmetry of G. The set of all automorphisms of G, denoted $\operatorname{Aut} G$, forms a group.

Conjugacy classes. Two group elements $a, b \in G$ are said to be conjugate to each other if there exists an element $g \in G$ such that $gag^{-1} = b$. In this case, we symbolically write $a \sim b$. Conjugation is an equivalence relation, since it is reflective ($a \sim a$), symmetric ($a \sim b$ iff $b \sim a$) and transitive (if $a \sim b$ and $b \sim c$ then $a \sim c$). Such a relation implies that G can be split into disjoint subsets $[a] \subset G$, called *conjugacy classes*, each containing all elements that are conjugate to each other:

$$[a] = \left\{ b \in G \mid gag^{-1} = b \;\; \text{for some} \;\; g \in G \right\} \ . \tag{1.2}$$

Obviously, a conjugacy class can be represented by any one of its elements, i.e. $[a] = [b]$ for all $b \sim a$. The number of distinct conjugacy classes is referred to as the *class number* of G, denoted here as $\mathrm{Cl}(G)$. All elements of a class have the same order. It is easy to see that an element constitutes a conjugacy class of its own if it commutes with all other elements of the group. As a result, in an Abelian group each class contains only one element and the class number equals the order of the group.

A common notation for conjugacy classes is writing the order of its elements, followed by an alphabetical letter. For example, $4A$ denotes a class of order four, $4B$ a different class of order four, $6A$ a class of order six, and so on. The identity is always a class of its own, namely $1A$, the unique class of order one.

Subgroups. A subgroup H is a subset $H \subset G$ which is itself a group, with the group structure inherited from G. Note that the identity element e always forms a subgroup $\{e\}$ of order 1, called the trivial subgroup. Subgroups H other than the trivial subgroup and G itself are called *proper subgroups* of G, and the notation $H < G$ is used for them (we use the notation $H \leq G$ if we can have $H = G$).

A *normal subgroup* N, also denoted as $N \triangleleft G$, is a subgroup of G that is invariant under conjugation by all elements of G:

$$N \triangleleft G \;\; \Leftrightarrow \;\; gNg^{-1} = N \text{ for all } g \in G \ . \tag{1.3}$$

As such, N is necessarily a union of conjugacy classes. A *maximal normal subgroup* of G is a normal subgroup which is not contained in any other normal subgroup of G, apart from G itself. Normal subgroups play a prominent role in quotient groups and group extensions (see below).

The centre $Z(G)$ of a group G is the set of all elements that commute with every other element, i.e.

$$Z(G) := \{a \in G \,|\, ab = ba \;\; \text{for all } b \in G\} \ . \tag{1.4}$$

The centre is always a normal subgroup of G. The *centralizer* of an element $g \in G$ is similarly defined by

$$C_G(g) = \{a \in G \,|\, ag = ga\}, \tag{1.5}$$

being the set of all elements that commute with g. Clearly, the centralizer of an element is always a subgroup of G.

Cosets. Let H be a subgroup of G, and take $g \in G$. We define the *left coset* of H in G with respect to g as the subset

$$gH = \{gh \mid h \in H\} \ . \tag{1.6}$$

The set of all left cosets of H in G is denoted by $G/H := \{gH \mid g \in G\}$. Similarly, the *right coset* of H in G with respect to g is defined as

$$Hg := \{hg \mid h \in H\} \ , \tag{1.7}$$

and the set of all right cosets of H in G is denoted by $H\backslash G := \{Hg \mid g \in G\}$. Notice that H is both a left and a right coset of itself, with respect to the identity $e \in G$.

One can more intuitively define left cosets in terms of an equivalence relation on G (not to be confused with conjugation); namely, for $a, b \in G$ we set $a \sim b$ iff $ah = b$ for some $h \in H$, i.e. a and b are related by multiplication of an element in H to the right. Then a, b represent the same equivalence class, which is exactly the coset $aH = bH$. All such classes make up G/H, which is viewed as a disjoint partition of G, as a set. Corresponding statements hold for right cosets. Some further facts about cosets:

- The number of left cosets is always equal to the number of right cosets, and is known as the *index* of H in G, denoted by $[G : H]$.
- If G is a finite group, then Lagrange's theorem states that the index equals the quotient of the order of G over the order of H, i.e. $[G : H]|H| = |G|$. This is indicatory of how G is partitioned under the coset equivalence relation associated with H.
- The left and the right cosets of H have the same number of elements, which is equal to the order of H.
- The left and right cosets of a normal subgroup coincide, as can be easily seen from its definition.

Quotient groups and group extensions. Cosets, like conjugacy classes, are in general not subgroups. However, given a normal subgroup N, the set G/N of right cosets (which coincides with the set of left cosets) inherits the group structure of G, and is called the *quotient group*. This can be seen by $(aN)(bN) = (ab)N$. The normal subgroup N can then be viewed as the kernel of the homomorphism $\psi : G \to G/N$. Note that in general G/N is not isomorphic to any subgroup of G. Moreover, the order of G/N is equal to the index $[G : N] = |G|/|N|$.

Consider now a *short exact sequence* of groups

$$1 \to M \xrightarrow{\phi} G \xrightarrow{\psi} Q \to 1 \ . \tag{1.8}$$

This means that $\phi(M)$, the embedding of M inside G, by ϕ is the kernel of the homomorphism ψ; in other words M is isomorphic to a normal subgroup $N \lhd G$, and $Q \cong G/N$. We then say that G is an *extension* of Q by M. An extension, as well as the corresponding sequence, is called *split* if there exists a homomorphism (embedding) $\widetilde{\psi} : Q \to G$ such that $\psi \circ \widetilde{\psi} = \mathrm{id}_Q$, the identity map on Q. We use the semi-direct product to denote such a split extension, $G = N \rtimes Q$. Otherwise, the extension is called *non-split*, and we write $G = N.Q$.

1.2 Classification of finite groups

From now on we focus on finite groups, which are groups with finite number of elements. The problem of classifying such groups can be reduced to the classification of a finite simple groups. A group is said to be *simple* if it has no proper normal subgroups. If G is not simple, then it can always be "decomposed" into a series of smaller groups, by considering quotients by maximal normal subgroups. To be precise, one can consider the *composition series*, which has the form

$$1 \lhd N_1 \lhd N_2 \lhd \cdots \lhd N_{n-1} \lhd N_n = G \ . \tag{1.9}$$

Here 1 denotes the trivial group, and every step of the series involves a maximal normal subgroup N_{i-1} of N_i, as well as the implied quotient group N_i/N_{i-1}. It can be shown that all the resulting quotient groups are simple, and the Jordan–Hölder theorem guarantees that for given G, two different composition series lead to the same simple groups. As a result, studying finite simple groups is to a large extent sufficient to understand general finite groups.

After a heroic effort spanning over half a century and involving more than 100 mathematicians leading to tens of thousands of pages of proof, all finite simple groups have been classified (see [5,6] for historical remarks). They belong to one of the following four categories: cyclic groups $\mathbb{Z}_p$ for prime p, alternating groups $\mathcal{A}_n$ ($n \geq 5$), 16 families of Lie type and 26 *sporadic groups*. Unlike the rest of finite simple groups, the 26 sporadic groups appear "sporadically" and are not part of infinite families. We will say more about the sporadic groups in the following section.

1.3 Sporadic groups and lattices

The largest sporadic group is the Fischer–Griess Monster group $\mathbb{M}$, which gets its name from its enormous size. The number of its element is

$$|\mathbb{M}| = 2^{46} \cdot 3^{20} \cdot 5^{9} \cdot 7^{6} \cdot 11^{2} \cdot 13^{3} \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \times 10^{53},$$

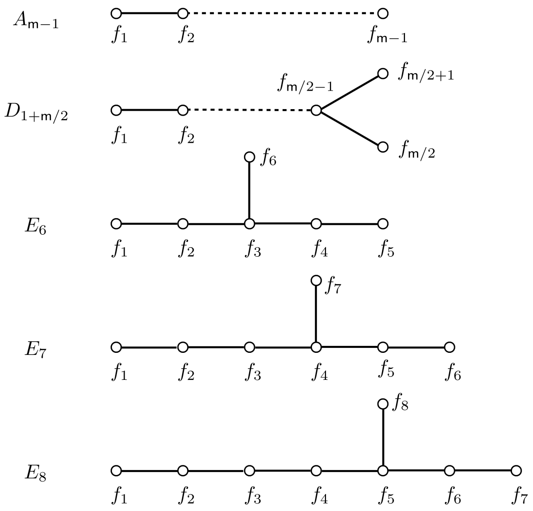
roughly the same as the number of atoms in the solar system. The Monster contains 20 of the 26 sporadic groups as its subgroups or quotients of subgroups, and these 20 is said to form 3 generations of a *happy family* by Robert Griess. In particular, the happy family includes the five Mathieu groups $M_{11}, M_{12}, M_{22}, M_{23}, M_{24}$. They are all subgroups of M_{24}, which is in turn a subgroup of the permutation group S_{24}, and are the the first sporadic groups to be discovered. The rest 6 which are not involved in the Monster are called the *pariahs* of sporadic groups.

The sporadic nature of the sporadic groups makes their existence somewhat mysterious and one might wonder what their "natural" representations are. An important hint is that many of the sporadic groups, especially those connected to the Monster, arise as subgroups of quotients of the automorphic groups of various special lattices. The appearance of moonshine involving sporadic groups sheds important light on the question, and the construction of moonshine often relies on the existence of these special lattices. As a result, in what follows we will briefly review the definition of lattices and their root systems, and introduce the special lattices relevant for moonshine.

Let V be a finite-dimensional real vector space of dimension r, equipped with an inner product $\langle \cdot, \cdot \rangle$. A finite subset $X \subset V$ of non-zero vectors is called a *root system of rank* r, if the following conditions are satisfied

- X spans V
- X is closed under reflections. Namely, $\beta - 2\dfrac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha \in X$ for all $\alpha, \beta \in X$.

Figure 1: The ADE Dynkin diagrams.



- The only multiples of $\alpha \in X$ that belong to X are α and $-\alpha$.
- For all $\alpha, \beta \in X$, we have $\dfrac{2\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$.

The elements $\alpha \in X$ of a root system are called *roots*. A root system X is said to be irreducible if it cannot be partitioned into proper orthogonal subsets $X = X_1 \cup X_2$. It turns out that the roots of such a system can have at most two possible lengths. If all roots have the same length, then the irreducible root system is called *simply-laced*. One can choose a subset Φ of roots $f_i \in X$ with $i = 1, \ldots, r$, such that each root can be written as an integral combination of $f_i \in \Phi$ with either all negative or all positive coefficients. Such a subset is called a set of *simple roots*, and is unique up to the action of the group generated by reflections with respect to all roots, called the *Weyl group* of X and denoted by $\mathrm{Weyl}(X)$.

To each irreducible root system we can attach a connected *Dynkin diagram*. Each simple root is associated with a node, and nodes associated to two distinct simple roots f_i, f_j are connected with N_{ij} lines, where

$$N_{ij} = \frac{2\langle f_i, f_j\rangle}{\langle f_i, f_i\rangle}\frac{2\langle f_j, f_i\rangle}{\langle f_j, f_j\rangle} \in \{0,1,2,3\}\ . \tag{1.10}$$

For simply-laced root systems we only have $N_{ij} \in \{0,1\}$. These correspond to the Dynkin diagrams of type A_n, D_n, E_6, E_7, E_8 with the subscript denoting the rank of the associated root system, as shown in figure 1.

Each irreducible root system contains a unique *highest root* θ with respect to a given set Φ of simple roots, whose decomposition

$$\theta = \sum_{i=1}^{r} a_i f_i \tag{1.11}$$

maximizes the sum $\sum a_i$. The *Coxeter number* of X is then defined by

$$m := 1 + \sum_{i=1}^{r} a_i\ . \tag{1.12}$$

The Coxeter number can also be defined in terms of $\mathrm{Weyl}(X)$. The product of reflections with respect to all simple roots $w = r_{f_1} r_{f_2} \cdots r_{f_r} \in W^X$ is called *Coxeter element*, and its order equals the Coxeter number m.

A *lattice* L of rank n is a free Abelian group isomorphic to the additive group $\mathbb{Z}^n$, equipped with a symmetric bilinear form $\langle \cdot, \cdot \rangle$. Embedding L into $\mathbb{R}^n$ gives the picture of a set of points inside the vector space $\mathbb{R}^n$. A few properties some lattices have that will be useful for us include the following:

- *Positive-definite*: the bilinear form induces a positive-definite inner product on $\mathbb{R}^n$.
- *Integral*: $\langle \lambda, \mu \rangle \in \mathbb{Z}$ for all $\lambda, \mu \in L$.
- *Even*: $\langle \lambda, \lambda \rangle \in 2\mathbb{Z}$ for all $\lambda \in L$.
- *Unimodular*: the dual lattice, defined by $L^* := \{\lambda \in L \otimes_{\mathbb{Z}} \mathbb{R} | \ \langle \lambda, L \rangle \subset \mathbb{Z}\}$, is isomorphic to the lattice itself.

All elements $\lambda \in L$ such that $\langle \lambda, \lambda \rangle = 2$ are called the *roots* of L.

Even, unimodular, positive-definite lattices in 24 dimensions play a distinguished role in several instances of moonshine, as we will discuss in the Part II and Part III of this note. It was proven by H. V. Niemeier in 1973 that there are only 24 inequivalent such lattices [7]. One of them, first discovered by J. Leech in 1967 and named the *Leech lattice*, is the only one of the 24 that has no root vectors [8–10]. The other 23, which we refer to as the *Niemeier lattices*, have non-trivial root systems. In fact, one useful construction of the Niemeier lattices is by combining the root lattices with the appropriate "glue vectors" [11]. It turns out that the 23 Niemeier lattices are uniquely labelled by the root systems X, called the *Niemeier root systems*, which are precisely one of the 23 unions of simply-laced (ADE) root systems $X = \cup_i Y_i$ satisfying the following conditions: 1) All components have the same Coxeter number, $\mathrm{Cox}(Y_i) = \mathrm{Cox}(Y_j)$; 2) the total rank equals the rank of the lattice $\sum_i \mathrm{rk}(Y_i) = 24$. Some examples out of the 23 include $24A_1$, $2A_{12}$, $8E_8$ and $D_{16}E_8$.

For each of these 24 even, unimodular, positive-definite lattices of rank 24 N we define a finite group

$$G_N := \mathrm{Aut}(N)/\mathrm{Weyl}(N), \tag{1.13}$$

where $\mathrm{Weyl}(N)$ denotes the Weyl group of the root system of N. In particular, when $N = \Lambda$ is the Leech lattice, the Weyl group $\mathrm{Weyl}(\Lambda)$ is the trivial group and $G_\Lambda \cong Co_0$ is the Conway group Co_0. By considering the quotients of this group and subgroups stabilising various structures we can obtain many of the sporadic groups. For instance, the sporadic simple group Co_1 is given by the quotient by the centre $Co_1 \cong Co_0/\{\pm 1\}$, and the Mathieu group M_{23} arises as the subgroup fixing a specific rank-2 sublattice. See Chapter 10 of [11] for a detailed discussion. If instead we choose N to be the Niemeier lattice with root system $24A_1$, for instance, the finite group $G_N \cong M_{24}$ is given by the largest Mathieu group. For the Niemeier lattice with root system $12A_2$, the finite group is $2.M_{12}$, the non-trivial extension of the Mathieu group M_{12}. These groups will play an important role in moonshine (cf. §4.2, §8.1, §5.2) and in the discussion of their physical context (cf. §8).

1.4 Representations of finite groups

The structure of groups is just what we need to describe symmetries. To put this into use we need the concept of representations of groups. In what follows we limit our discussion to *complex representations*, arising from group acting on a complex vector space V. More precisely, consider the group homomorphism $\rho : G \to GL(V)$. We can think of $\rho(g)$ as invertible $n \times n$ complex matrices. In particular we have $\rho\left(g^{-1}\right) = (\rho(g))^{-1}$. The vector space together with the map (V, ρ) is called a *representation* of dimension n. Often one refers to either V or ρ as the representation, while implicitly referring to the full data. The vector space V is also called a G-module in this context, and is said to carry a G-action. We say that the G-action is *faithful*, if no two distinct elements $g, g' \in G$ lead to $\rho(g) = \rho(g')$ (the corresponding representation is also called faithful).

Irreducible representations and dual representations. Given two representations (V, ρ) and (V', ρ') one can define their direct sum and their tensor product in a straightforward way, which leads to new representations $V \oplus V'$ and $V \otimes V'$.

Two representations ρ, ρ' are *equivalent* if there exists an invertible $n \times n$ matrix M such that $M\rho'(g) = \rho(g)M$ for all $g \in G$. A *subrepresentation* of a representation (V, ρ) is a representation (U, ρ'), where U is a subspace $U \subset V$ that is preserved by the action of G, and ρ' is the restriction of ρ to U. A representation V is said to be *irreducible* if it does not contain any proper subrepresentation, and *indecomposable* if it cannot be written as a direct sum of two (or more) non-zero subrepresentations. For finite groups, these two notions coincide. A representation is called *completely reducible* if it is a direct sum of finitely many irreducible representations, i.e. if it can be fully decomposed into irreducible pieces.

The so-called Maschke's theorem states that all (finite-dimensional) representations of a finite group are always completely reducible. There are two steps for proving this. First we show that a unitary representation is always completely reducible, by using the fact that given an inner product $\{\cdot, \cdot\} : V \times V \to \mathbb{C}$, the orthogonal complement of U in V is also a subrepresentation if U itself is a subrepresentation of V. Next we show that any representation is unitary with respect to the group-invariant inner product

$$\{v, w\} := \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)v, \rho(g)w \rangle, \quad v, w \in V \ , \tag{1.14}$$

which then completes the proof.

We also mention the *dual representation* ρ^* of a representation ρ, defined by

$$\rho^*(g) := \left(\rho\left(g^{-1}\right)\right)^T \ , \quad g \in G \ , \tag{1.15}$$

which is the natural group action on the dual space $V^* = \mathrm{End}(V, \mathbb{C})$. Taking ρ to be unitary, we have $\rho^*(g) = \overline{\rho(g)}$. In other words, the dual representation is equivalent to the complex conjugate representation.

Characters. The character χ_ρ of a representation (V, ρ), with V a vector space over $\mathbb{C}$, is a map $G \to \mathbb{C}$ defined by the trace of the representation matrices,

$$\chi_\rho(g) := \mathrm{Tr}(\rho(g)) \ , \quad g \in G. \tag{1.16}$$

We will also often denote this trace by $\mathrm{Tr}_V\, g$. If ρ is irreducible, χ_ρ is called an irreducible character. Some properties of characters (for finite groups) are summarized below:

- The character is a *class function*, i.e. $\chi_\rho\left(hgh^{-1}\right) = \chi_\rho(g) \quad \forall\ g, h \in G$. This follows directly from the cyclic property of the trace.
- Two complex representations for a finite group have the same characters if and only if they are equivalent, which can be shown using the orthogonality property discussed below.
- The restriction of a character of G to a subgroup $H < G$ is a character of H.
- $\chi_\rho\left(g^{-1}\right) = \overline{\chi_\rho(g)}$, as follows from the fact that all eigenvalues of $\rho(g)$ are $|g|$-th roots of unity.
- For be two representations ρ, ρ' of G and $g \in G$, we have:

  $$\chi_{\rho\oplus\rho'}(g) = \chi_\rho(g) + \chi_{\rho'}(g)\ ,\quad \chi_{\rho\otimes\rho'}(g) = \chi_\rho(g)\chi_{\rho'}(g)\ ,\quad \chi_{\rho^*}(g) = \overline{\chi_\rho(g)}\ . \tag{1.17}$$

  This means that the characters form a commutative and associative algebra.

Orthogonality. Due to Schur's orthogonality relations (e.g. §4 of [2]), characters of unitary representations are equipped with a Hermitian inner product,

$$\langle \chi_\rho, \chi_{\rho'} \rangle = \frac{1}{|G|} \sum_{g\in G} \chi_\rho(g)\overline{\chi_{\rho'}(g)}\ . \tag{1.18}$$

When ρ and ρ' are irreducible representations, one can show that $\langle \chi_\rho, \chi_{\rho'} \rangle = 1$ if the two irreducible representations are equivalent, and it vanishes otherwise. As a result, characters of irreducible representations are orthonormal vectors in the space of class functions. In fact, it is possible to show that they span this space, from which one can conclude the important fact that the number of (inequivalent) irreducible representations equals the number of conjugacy classes (see for instance §3-7 of [1]). Moreover, one can show that there is another orthonormality property,

$$\sum_\rho \chi_\rho(g)\ \overline{\chi_\rho(h)} = \begin{cases} |C_G(g)|, & h \in [g] \\ 0, & \text{otherwise} \end{cases}\ , \tag{1.19}$$

where the sum is over inequivalent irreducible representations, and $|C_G(g)|$ denotes the order of the centralizer of $g \in G$, which is equal to the order of the group divided by the number of elements in the conjugacy class $[g]$.

Character table. We have already mentioned that the number of irreducible representations of a finite group G is equal to the number of conjugacy classes of G. We can group all characters of G into its *character table*, which is a square table of size $\mathrm{Cl}(G) \times \mathrm{Cl}(G)$, with rows labelling the different irreducible representations and columns labelling the different conjugacy classes. In other words, the (i, j) component of the character table is the character $\chi_i(g)$ of the i-th irreducible representation, evaluated at any g in the j-th conjugacy class . As an example, the character table for the alternating group $\mathcal{A}_6$ is displayed in Table 1. Note that there is an additional piece of information in the above table, the so-called power map. The row starting with sP gives the conjugacy classes $[g^s]$.

		1A	2A	3A	3B	4A	5A	5B
	2P	1A	1A	3A	3B	2A	5B	5A
	3P	1A	2A	1A	1A	4A	5B	5A
	5P	1A	2A	3A	3B	4A	1A	1A
χ_1		1	1	1	1	1	1	1
χ_2		5	1	2	-1	-1	0	0
χ_3		5	1	-1	2	-1	0	0
χ_4		8	0	-1	-1	0	A	${}^*A$
χ_5		8	0	-1	-1	0	${}^*A$	A
χ_6		9	1	0	0	1	-1	-1
χ_7		10	-2	1	1	0	0	0

Table 1: Character table for $\mathcal{A}_6$, where $A = \frac{1-\sqrt{5}}{2}$ and ${}^*A = \frac{1+\sqrt{5}}{2}$.

Supermodules. We say that a G-module on a superspace ($\mathbb{Z}_2$-graded vector space) is a G-supermodule. Explicitly, if V is a G-supermodule it has the structure

$$V = V_{\underline{0}} \oplus V_{\underline{1}} \tag{1.20}$$

where $V_{\underline{0}}$ and $V_{\underline{1}}$ are both G-modules. We will sometimes refer to V as a *virtual representation* of G. The supertrace Str is defined to act with a minus sign on the odd subspaces: $\mathrm{Str}_V g := \mathrm{Tr}_{V_{\underline{0}}} g - \mathrm{Tr}_{V_{\underline{1}}} g$.

Cycle shapes and Frame shapes. As the name suggests, an N-dimensional *permutation representation* ρ_p of a group G has as its representation matrices $N \times N$ permutation matrices (all elements zero, apart from a single entry of 1 in each row and column). Given such a representation, to each conjugacy class in such a representation we can associate a *cycle shape*, which encodes the number and type of permutation cycles that elements of this class correspond to. A cycle shape has the general form

$$n_1^{\ell_1} n_2^{\ell_2} \cdots n_r^{\ell_r} \ , \quad \sum_{s=1}^{r} \ell_s n_s = N \ , \tag{1.21}$$

where n_s, ℓ_s are all positive integers, and n denotes an *n-cycle*, i.e. it represents a permutation of n elements. The exponents ℓ_s count the number of n_s-cycles. Clearly, an order k element can only have cycles of size n_s which divides k. Note that the cycle shapes can be read directly off the character table, including the power map.

More generally, we can define the *Frame shape* of $g \in G$ given any representation ρ, provided that all characters of ρ are rational numbers. Their rationality ensures that if λ is an eigenvalue of g, then λ^k is also an eigenvalue when k is co-prime to $|g|$ [12]. Denoting by $\lambda_1, \lambda_2, \ldots, \lambda_N$ the g-eigenvalues, then there exists a set of positive integers $i_1, i_2, \ldots, i_s$ and a set of non-zero integers $\ell_1, \ell_2, \ldots, \ell_s$ with the same cardinality such that

$$\det(\mathbf{1} - t\rho(g)) = \prod_{i=1}^{N} (1 - t\lambda_i) = \prod_{r=1}^{s} (1 - t^{i_r})^{\ell_r} \ . \tag{1.22}$$

Clearly one must have $\sum_{s=1}^{r} \ell_s i_s = N$ and we call $i_1^{\ell_1} i_2^{\ell_2} \cdots i_r^{\ell_r}$ the Frame shape of the conjugacy class $[g]$ for the representation ρ.

2 Modular objects

In this chapter we will briefly introduce the concept of modular forms and their extensions including mock modular forms, Jacobi forms, and mock Jacobi forms.

2.1 Modular forms

One of the standard references on modular forms, which we partially follow here, is [13]. It is well-known that $SL_2(\mathbb{R})$ acts on the upper-half plane $\mathbb{H} := \{\tau \in \mathbb{C} | \ \Im(\tau) > 0\}$ by a fractional linear (Möbius) transformation:

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \mathbb{H} \to \mathbb{H} \ , \quad \tau \mapsto \gamma\tau := \frac{a\tau + b}{c\tau + d} \ . \tag{2.1}$$

In defining modular forms we consider discrete subgroups of $SL_2(\mathbb{R})$, an important example of which is the modular group $SL_2(\mathbb{Z})$. It is generated by

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.2}$$

satisfying $(ST)^3 = 1$ and $S^2 = -1$. We will also often work with $PSL_2(\mathbb{Z}) \cong SL_2(\mathbb{Z})/\{\pm 1\}$, which is also the mapping class group of the torus (cf. §3). We will often also consider the upper-half plane extended by adding the *cusps* $\{i\infty\} \cup \mathbb{Q}$, which $SL_2(\mathbb{Z})$ acts transitively on as one can see from $\gamma\infty = \frac{a}{c}$.

We start by defining weight zero modular forms on the modular group $SL_2(\mathbb{Z})$, which are simply holomorphic functions on $\mathbb{H}$ that are invariant under the action of $SL_2(\mathbb{Z})$:

$$f(\tau) = f(\gamma\tau) \quad \forall\, \gamma \in SL_2(\mathbb{Z}). \tag{2.3}$$

In particular, f has to be holomorphic as τ approaches the boundary of $\mathbb{H}$ at the cusps $\{i\infty\} \cup \mathbb{Q}$. But this turns out to be too restrictive: basic complex analysis tells us that constants are the only such functions. As a result, we would like to further generalise the above definition in the following directions:

1. *Analyticity*: the function is allowed to have exponential growth near the cusps. Such functions are said to be *weakly holomorphic modular forms.*

2. *Weights*: one allows for a scaling factor in the transformation rule. See (2.6).

3. *Other Groups*: one replaces $SL_2(\mathbb{Z})$ by a general $\Gamma < SL_2(\mathbb{R})$ in the transformation property (2.3).

4. *Multipliers*: one modifies the transformation rule (2.3) by allowing for a non-trivial character $\psi : SL_2(\mathbb{Z}) \to \mathbb{C}^*$. See (2.6).

5. *Vector-Valued*: instead of $f : \mathbb{H} \to \mathbb{C}$ we consider a vector-valued function $f : \mathbb{H} \to \mathbb{C}^n$ with n components.

Of course, the above generalisations can be combined. For instance one can consider a vector-valued modular form with multipliers for a subgroup Γ of $SL_2(\mathbb{R})$. Obviously, in the vector-valued case the character ψ is no longer a phase but a matrix. Also, the above concepts are not entirely independent. For instance, a component of a vector-valued modular form for $SL_2(\mathbb{Z})$ can be considered as a (single-valued) modular form for a subgroup of $SL_2(\mathbb{Z})$, and vice versa.

We will first start with the first generalisation and introduce the concept of *modular functions*. We say that $f : \mathbb{H} \to \mathbb{C}$ is a modular function if f is meromorphic in $\mathbb{H}$, satisfies the transformation rule (2.3), and grows like $e^{2\pi i \tau m}$ for some $m > -\infty$. In fact, modular functions form a function field with a single generator, called the *Hauptmodul* or *principal modulus*. This is because the fundamental domain $SL_2(\mathbb{Z})\backslash\mathbb{H}$ is a genus zero Riemann surface when finitely many points are added. Writing the upper-half plane with the cusps attached as $\hat{\mathbb{H}} = \mathbb{H} \cup \{i\infty\} \cup \mathbb{Q}$, the Hauptmodul has the property that it is an isomorphism between the two spheres $SL_2(\mathbb{Z})\backslash\hat{\mathbb{H}}$ and $\hat{\mathbb{C}}$. Clearly, such a Hauptmodul is unique up to Möbius transformations, or the choice of three points on the sphere. As a result, there is a unique Hauptmodul with the expansion

$$J(\tau) = q^{-1} + O(q) \tag{2.4}$$

near $\tau \to i\infty$. Here and in what follows we will write $q := \mathrm{e}(\tau)$, where $\mathrm{e}(x) := e^{2\pi i x}$ for $x \in \mathbb{C}$. In terms of the Eisenstein series and Dedekind eta function (cf. (2.7) and (2.19)), the J-function is given by

$$J(\tau) = j(\tau) - 744 = \frac{E_4^3(\tau)}{\eta^{24}(\tau)} - 744. \tag{2.5}$$

In general, a Hauptmodul can be defined as the generator of the field of modular functions for $\Gamma \le SL_2(\mathbb{R})$ whenever $\Gamma\backslash\hat{\mathbb{H}}$ is genus zero. These Hauptmoduls play an important role in moonshine.

Apart from the definition given above, there are three other equivalent ways of viewing modular functions. First, due to (2.3) we can view f as a function from the suitably compactified fundamental domain $SL_2(\mathbb{Z})\backslash\mathbb{H}$ to the Riemann sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Second, due to the relation between $SL_2(\mathbb{Z})$ and rank two lattices we can associate to each τ a complex lattice $\Lambda_\tau := 1 \cdot \mathbb{Z} + \tau \cdot \mathbb{Z}$, and identify f as a function that associates to each such lattice Λ_τ a complex function $f(\tau)$, which is moreover invariant under a rescaling $\Lambda \mapsto \lambda\Lambda$, $\lambda \in \mathbb{C}$ of the lattice. The third way, which plays an important role in the a relation between modular forms and 2-dimensional conformal field theories, stems from the interpretation of $SL_2(\mathbb{Z})\backslash\mathbb{H}$ as the complex structure moduli space of a Riemann surface of genus one. This can be easily understood from the fact that a torus can be described as the complex plane modulo a rank two lattice, and is therefore up to a scale given by $\mathbb{C}/\Lambda_\tau$ for some $\tau \in \mathbb{H}$. The modular function can then be thought of as associating to each torus a complex number which only depends on its complex structure modulus τ. In this context, the group $PSL_2(\mathbb{Z}) := SL_2(\mathbb{Z})/\{\mathbb{1}, -\mathbb{1}\}$ plays the role of the mapping class group of a torus (cf. §3), where the $\mathbb{Z}_2 = \{\mathbb{1}, -\mathbb{1}\}$ central subgroup acts trivially on $\mathbb{H}$.

Next we turn to the second generalisation and introduce modular forms on the modular group $SL_2(\mathbb{Z})$ of a general weight k. They are defined as holomorphic functions on $\mathbb{H}$ that transform under the action of $SL_2(\mathbb{Z})$ as:

$$f(\tau) = (c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right) \quad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \,. \tag{2.6}$$

From the lattice point of view, we consider complex functions f associated to a lattice Λ that scale like $f \mapsto \lambda^{-k} f$ under a rescaling $\Lambda \mapsto \lambda\Lambda$ of the lattice. We will consider integral and half-integral weight k.[1]

With this definition we start to get some non-trivial examples, even when holomorphicity at the cusps is required. For instance the following Eisenstein series:

$$\begin{aligned} E_4(\tau) &= 1 + 240 \sum_{n=1}^{\infty} \frac{n^3 q^n}{1-q^n} = 1 + 240\, q + 2160\, q^2 + \dots \ , \\ E_6(\tau) &= 1 - 504 \sum_{n=1}^{\infty} \frac{n^5 q^n}{1-q^n} = 1 - 504\, q - 16632\, q^2 + \dots \ , \end{aligned} \tag{2.7}$$

are examples of modular forms of weight 4 and weight 6, respectively. But the definition is still somewhat too restrictive as these two Eisenstein series are all there is: the ring of modular forms on $SL_2(\mathbb{Z})$ is generated freely by E_4 and E_6. Namely, any modular form of integral weight k can be written (uniquely) as a sum of monomials $E_4^{\alpha} E_6^{\beta}$ with $k = 4\alpha + 6\beta$. We denote the space of modular forms of weight k for group Γ by $M_k(\Gamma)$. Among modular forms, the so-called cusp forms are often of special interest. We say that a modular form f of weight k is a cusp form if $y^{k/2} f(x+iy)$ is bounded as $y \to \infty$. This condition guarantees that f has vanishing constants in its Fourier coefficients at all cusps.

In the third type of generalisation, we often encounter the $SL_2(\mathbb{Z})$ subgroups defined by the following congruences. For a positive integer N, we define

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \ \mid c \equiv 0 \bmod N \right\} \ . \tag{2.8}$$

Below we will illustrate the generalisations above with some examples.

First we consider the Jacobi theta functions. Consider a 1-dimensional lattice with bilinear form $\langle x, x \rangle = x^2$. The associated theta function is

$$\theta_3(\tau) = \sum_{n \in \mathbb{Z}} q^{n^2/2} \, . \tag{2.9}$$

This simple function turns out to admit an expression in terms of infinite products

$$\theta_3(\tau) = \sum_{n \in \mathbb{Z}} q^{n^2/2} = \prod_{n=1}^{\infty} (1-q^n)(1+q^{n-1/2})^2 \ , \tag{2.10}$$

and has nice modular properties. To describe the modular properties, it is most natural to introduce another two theta functions,

$$\theta_2(\tau) = \sum_{n+\frac{1}{2} \in \mathbb{Z}} q^{n^2/2} = 2 q^{1/8} \prod_{n=1}^{\infty} (1-q^n)(1+q^n)^2 \ , \tag{2.11}$$

$$\theta_4(\tau) = \sum_{n \in \mathbb{Z}} (-1)^n q^{n^2/2} = \prod_{n=1}^{\infty} (1-q^n)(1-q^{n-1/2})^2 \ . \tag{2.12}$$

It turns out that they are the three components of a vector-valued modular form for $SL_2(\mathbb{Z})$

$$\Theta(\tau) := \begin{pmatrix} \theta_2(\tau) \\ \theta_3(\tau) \\ \theta_4(\tau) \end{pmatrix} , \quad \Theta(\tau) = \sqrt{\frac{i}{\tau}} \mathcal{S}\, \Theta\left(-\frac{1}{\tau}\right) = \mathcal{T}\, \Theta(\tau+1) \, , \tag{2.13}$$

---

[1] Clearly, special care needs to be taken when k is half-integral. Strictly speaking, one should work with the metaplectic double cover of $SL_2(\mathbb{Z})$. However we will avoid discussing the subtleties here as they will not cause any difficulty for us. We will refer the reader to [14] for more details.

where

$$\mathcal{S} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathcal{T} = \begin{pmatrix} e(-\frac{1}{8}) & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.14}$$

To illustrate the relation between vector-valued modular forms and modular forms for a congruence subgroup, consider $\theta(\tau) := \theta_3(2\tau)$. This transforms in the following way as a weight $1/2$ modular form for $\Gamma_0(4)$ with a non-trivial multiplier:

$$\theta(\tau) = \left(\frac{c}{d}\right) \epsilon_d \; (c\tau + d)^{-\frac{1}{2}} \; \theta(\gamma\tau) \tag{2.15}$$

for all $\gamma \in \Gamma_0(4)$, where

$$\epsilon_d := \begin{cases} 1 \, , & d \equiv 1 \bmod 4 \\ i \, , & d \equiv 3 \bmod 4 \end{cases}$$

and the Legendre symbol used above is defined as[2]

$$\left(\frac{\kappa}{\lambda}\right) := \begin{cases} +1 \, , & \text{if } \; \kappa \neq 0 \bmod \lambda \text{ and } \kappa \text{ is a quadratic residue modulo } \lambda \\ -1 \, , & \text{if } \; \kappa \neq 0 \bmod \lambda \text{ and } \kappa \text{ is not a quadratic residue modulo } \lambda \\ \;\; 0 \, , & \text{if } \; \kappa \equiv 0 \bmod \lambda \end{cases} .$$

Later we will see that these theta functions can be naturally considered as the specialisation at $z = 0$ of the two-variable Jacobi theta functions, defined either as infinite sums or infinite products:

$$\begin{aligned}
\theta_1(\tau, z) &= -i \sum_{n+\frac{1}{2}\in\mathbb{Z}} (-1)^{n-\frac{1}{2}} y^n q^{n^2/2} \\
&= -i q^{1/8} (y^{1/2} - y^{-1/2}) \prod_{n=1}^{\infty} (1-q^n)(1-yq^n)(1-y^{-1}q^n) \, , \\
\theta_2(\tau, z) &= \sum_{n+\frac{1}{2}\in\mathbb{Z}} y^n q^{n^2/2} \\
&= (y^{1/2} + y^{-1/2}) q^{1/8} \prod_{n=1}^{\infty} (1-q^n)(1+yq^n)(1+y^{-1}q^n) \, , \\
\theta_3(\tau, z) &= \sum_{n\in\mathbb{Z}} y^n q^{n^2/2} = \prod_{n=1}^{\infty} (1-q^n)(1+yq^{n-1/2})(1+y^{-1}q^{n-1/2}) \, , \\
\theta_4(\tau, z) &= \sum_{n\in\mathbb{Z}} (-1)^n y^n q^{n^2/2} = \prod_{n=1}^{\infty} (1-q^n)(1-yq^{n-1/2})(1-y^{-1}q^{n-1/2}) \quad .
\end{aligned}$$

They transform in the following way.

Let

$$\Theta(\tau, z) := \begin{pmatrix} \theta_1(\tau, z) \\ \theta_2(\tau, z) \\ \theta_3(\tau, z) \\ \theta_4(\tau, z) \end{pmatrix} \quad , \tag{2.16}$$

then we have (cf. §2.2),

$$\Theta(\tau, z) = \sqrt{\frac{i}{\tau}} \, e\big(-\frac{z^2}{2\tau}\big) \, \mathcal{S}' \, \Theta\big(-\frac{1}{\tau}, \frac{z}{\tau}\big) = \mathcal{T}' \, \Theta(\tau + 1, z) \, , \tag{2.17}$$

---

[2] κ is said to be a quadratic residue modulo λ if $\exists \; x \in \mathbb{Z}$ such that $x^2 \equiv \kappa \bmod \lambda$.

where

$$\mathcal{S}' = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathcal{T}' = \begin{pmatrix} e(-1/8) & 0 & 0 & 0 \\ 0 & e(-1/8) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.18}$$

Another modular form one frequently encounters is the *Dedekind eta function*

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n) \ . \tag{2.19}$$

It is a weight 1/2 modular form with a non-trivial multiplier, satisfying

$$\eta(\tau) = \sqrt{\frac{i}{\tau}} \eta\big(-\frac{1}{\tau}\big), \quad \eta(\tau) = e\big(-\frac{1}{24}\big)\, \eta(\tau+1). \tag{2.20}$$

It is related to the theta functions by

$$\eta(\tau)^3 = \frac{1}{2} \theta_2(\tau)\theta_3(\tau)\theta_4(\tau). \tag{2.21}$$

Its 24-th power $\Delta := \eta^{24}$ is the familiar weight 12 cusp form for the modular group $SL_2(\mathbb{Z})$.

2.2 (Skew-)Holomorphic Jacobi forms

In this subsection we collect the definitions of (skew-)holomorphic Jacobi forms. These types of objects play a crucial role in moonshine and its connection to physics. This subsection, consisting mostly of definitions, follows §3.1 of [15] very closely.

We first define elliptic forms [16]. For m an integer define the index m *elliptic action* of the group $\mathbb{Z}^2$ on functions $\phi : \mathbb{H} \times \mathbb{C} \to \mathbb{C}$ by setting

$$(\phi|_m(\lambda, \mu))(\tau, z) := \mathrm{e}(m\lambda^2\tau + 2m\lambda z)\, \phi(\tau, z + \lambda\tau + \mu) \tag{2.22}$$

for $(\lambda, \mu) \in \mathbb{Z}^2$. Say that a smooth function $\phi : \mathbb{H} \times \mathbb{C} \to \mathbb{C}$ is an *elliptic form* of index m if $z \mapsto \phi(\tau, z)$ is holomorphic. Denote by $\mathcal{E}_m$ the space of elliptic forms of index m. Observe that any elliptic form $\phi \in \mathcal{E}_m$ admits a *theta-decomposition*

$$\phi(\tau, z) = \sum_{r \bmod 2m} h_r(\tau)\theta_{m,r}(\tau, z) \ , \tag{2.23}$$

where the theta series are given by

$$\theta_{m,r}(\tau, z) := \sum_{\ell = r \bmod 2m} q^{\ell^2/4m} y^{\ell} \ ,$$

for some $2m$ smooth functions $h_r : \mathbb{H} \to \mathbb{C}$. To see this, note from $\phi(\tau, z) = \phi(\tau, z+1)$ that we have $\phi(\tau, z) = \sum_{\ell \in \mathbb{Z}} c_\ell(\tau) y^\ell$ for some $c_\ell : \mathbb{H} \to \mathbb{C}$. Then the identity $\phi|_m(1, 0) = \phi$ implies that $c_r(\tau)q^{-r^2/4m}$ depends only on $r \bmod 2m$. The $2m$ functions $h_r(\tau) := c_r(\tau)q^{-r^2/4m}$ are precisely the theta-coefficients of ϕ in the theta-decomposition.

It will be convenient to regard h_r and $\theta_{m,r}$ in (2.23) as defining $2m$-vector-valued functions $h := (h_r)_{r \bmod 2m}$ and $\theta_m := (\theta_{m,r})_{r \bmod 2m}$. Then the theta-decomposition (2.23) may be more succinctly written as $\phi = h^t\theta_m$, where the superscript t denotes matrix transpose.

It follows from the Poisson summation formula that the vector-valued function $\theta_m = (\theta_{m,r})$ has the following behaviour under $SL_2(\mathbb{Z})$:

$$\theta_m\left(-\frac{1}{\tau},\frac{z}{\tau}\right)\frac{1}{\sqrt{\tau}}\,\mathrm{e}\left(-\frac{mz^2}{\tau}\right) = \mathcal{S}\theta_m(\tau,z), \quad \theta_m(\tau+1,z) = \mathcal{T}\theta_m(\tau,z), \tag{2.24}$$

where $\mathcal{S} = (\mathcal{S}_{rr'})$ and $\mathcal{T} = (\mathcal{T}_{rr'})$ are unitary matrices defined for a fixed positive integer m, given by $\mathcal{S}_{rr'} := \frac{1}{\sqrt{2m}}\,\mathrm{e}\left(-\frac{1}{8}-\frac{rr'}{2m}\right)$ and $\mathcal{T}_{rr'} := \mathrm{e}\left(\frac{r^2}{4m}\right)\delta_{r,r'}$. (Cf. e.g. §5 of [17].) This suggests that we obtain elliptic forms $\phi = h^t\theta_m \in \mathcal{E}_m$ with good modular transformation properties $SL_2(\mathbb{Z})$ by requiring suitable conditions on h.

To formulate these notions precisely, define the weight k *modular*, and *skew-modular* actions of $SL_2(\mathbb{Z})$ on $\mathcal{E}_m$, for k and m integers, by setting

$$\begin{aligned}(\phi|_{k,m}\gamma)(\tau,z) &:= \phi\left(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\right)\frac{1}{(c\tau+d)^k}\,\mathrm{e}\left(-\frac{cmz^2}{c\tau+d}\right)\\ (\phi|^{\mathrm{sk}}_{k,m}\gamma)(\tau,z) &:= \phi\left(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\right)\frac{1}{(c\bar\tau+d)^k}\frac{c\bar\tau+d}{|c\tau+d|}\,\mathrm{e}\left(-\frac{cmz^2}{c\tau+d}\right)\ ,\end{aligned} \tag{2.25}$$

for $\phi \in \mathcal{E}_m$ and $\gamma = \left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in SL_2(\mathbb{Z})$.

Roughly speaking, a Jacobi form of weight k and index m is an elliptic form $\phi(\tau,z)$, holomorphic in the τ-variable, which is moreover invariant under $|_{k,m}\gamma$ for all $\gamma \in SL_2(\mathbb{Z})$. Note that $\phi|_{k,m}\left(\begin{smallmatrix} 1 & 1\\ 0 & 1\end{smallmatrix}\right) = \phi$ implies the expansion

$$\phi(\tau,z) = \sum_{\substack{D,\ell\in\mathbb{Z}\\ D=\ell^2 \bmod 4m}} C_\phi(D,\ell)q^{-D/4m}q^{\ell^2/4m}y^\ell\ , \tag{2.26}$$

where $C_\phi(D,\ell)$ depends only on $\ell \bmod 2m$, corresponding to the theta decomposition

$$h_r(\tau) = \sum_{\substack{D\in\mathbb{Z}\\ D=r^2 \bmod 4m}} C_\phi(D,r)q^{-D/4m}\ . \tag{2.27}$$

The invariance under $SL_2(\mathbb{Z})$ of $\phi = h^t\theta_m$ leads to the modularity of the vector-valued function $h = (h_r)$. In other words, $h = (h_r)$ transforms as a vector-valued modular form and contains precisely the same information as the Jacobi form. To complete the definition, we also need to specify the growth behaviour of $h(\tau)$ near the cusp. We say that $\phi \in \mathcal{E}_m$, invariant under $|_{k,m}\gamma$ for all $\gamma \in SL_2(\mathbb{Z})$, is a *weak holomorphic/holomorphic/cuspidal holomorphic Jacobi form* if the Fourier coefficients satisfy $C_\phi(D,r) = 0$ unless $-D + r'^2 \geq 0$ for all $r' = r \bmod 2m$, $C_\phi(D,r) = 0$ for $D > 0$, or $C_\phi(D,r) = 0$ for $D \geq 0$, respectively. We denote the space of weak holomorphic Jacobi forms of weight k and index m by $J^{\mathrm{wk}}_{k,m}$. Notice that, at odd weight, applying (2.25) to the case $\gamma = -\left(\begin{smallmatrix} 1 & 0\\ 0 & 1\end{smallmatrix}\right)$ shows that the Jacobi form must be odd under $z \leftrightarrow -z$. It will therefore be convenient to introduce

$$\hat\theta_{m,r} = \theta_{m,r} - \theta_{m,-r}\ . \tag{2.28}$$

We now turn to the closely related skew-holomorphic Jacobi forms. An elliptic form $\phi \in \mathcal{E}_m$ is called a *weak skew-holomorphic Jacobi form* if it meets the following conditions. First, its theta-coefficients are anti-holomorphic functions on $\mathbb{H}$; second, it is invariant for the weight k skew-modular action (2.25), so that $\phi|^{\mathrm{sk}}_{k,m}\gamma = \phi$ for all $\gamma \in SL_2(\mathbb{Z})$; finally, $\tau \mapsto \phi(\tau, z)$ remains bounded as $\Im(\tau) \to \infty$ for fixed $z \in \mathbb{C}$. Thus a weak skew-holomorphic Jacobi form admits a Fourier expansion of the form

$$\phi(\tau, z) = \sum_{\substack{D,\ell \in \mathbb{Z} \\ D=\ell^2 \bmod 4m}} C_\phi(D, \ell)\, \bar{q}^{D/4m} q^{\ell^2/4m} y^\ell \ , \tag{2.29}$$

for some $2m$ functions $D \mapsto C_\phi(D, r)$, and we recover its theta-coefficients by writing

$$h_r(\tau) = \sum_{\substack{D \in \mathbb{Z} \\ D=r^2 \bmod 4m}} C_\phi(D, r) \bar{q}^{D/4m} \ . \tag{2.30}$$

A weak skew-holomorphic Jacobi form ϕ is called a *skew-holomorphic Jacobi form*, or a *cuspidal skew-holomorphic Jacobi form*, when the Fourier coefficients satisfy $C_\phi(D, r) = 0$ for $D < 0$, or $C_\phi(D, r) = 0$ for $D \leq 0$, respectively.

We will close this subsection with some examples.

- Define

  $$\begin{aligned} \phi_{0,1}(\tau, z) &= 4 \sum_{i=2,3,4} \Big(\frac{\theta_i(\tau, z)}{\theta_i(\tau, 0)}\Big)^2 \ , \\ \phi_{-2,1} &= -\frac{\theta_1(\tau, z)^2}{\eta^6(\tau)} \ . \end{aligned} \tag{2.31}$$

  The ring of weak Jacobi forms of even weight is freely generated by $\phi_{0,1}$ and $\phi_{-2,1}$ over the ring of modular forms for $SL_2(\mathbb{Z})$:

  $$J^{\rm wk}_{2k,m} = \sum_{j=0}^{m} M_{2k+2j}(SL_2(\mathbb{Z}))\, \phi_{-2,1}^j \phi_{0,1}^{m-j} \ . \tag{2.32}$$

  The function $\phi_{0,1}$ plays an important role in Mathieu and umbral moonshine, since $2\phi_{0,1}$ coincides with the $K3$ elliptic genus $\mathbf{EG}(\tau, z; K3)$. See §3.3 for a definition of the elliptic genus.

- In [18] it was shown that the modified elliptic genus of the so-called MSW string [19] involves a skew-holomorphic Jacobi form with index specified by the charges of the black hole, provided that the moduli is fixed at their black hole attractor value. See §9 for more discussions on this.

  As a concrete example, consider a single M5 brane wrapping $\mathbb{P}^2$. The modified elliptic genus of the resulting effective string is given by $\frac{1}{2\eta^6(\tau)} \overline{t_{2,2}(\tau, \frac{\bar{z}}{2})}$, where

  $$t_{2,2}(\tau, z) = \sum_{r \bmod 4} \overline{\theta^1_{2,r}(\tau)} \theta_{2,r}(\tau, z)$$

  is a weight 2 skew-holomorphic Jacobi forms. In the above we have used the definition

  $$\theta^1_{m,r}(\tau) := \frac{1}{2\pi i} \frac{\partial}{\partial z} \theta_{m,r}(\tau, z)|_{z=0} \ . \tag{2.33}$$

  Note that $\theta^1_{2,1}(\tau) = -\theta^1_{2,-1}(\tau) = \eta^3(\tau)$ coincides with the shadow of the Mathieu moonshine function (2.42).

2.3 Mock modular forms

In this section we introduce mock modular forms and the closely related concept of mock Jacobi forms. We follow the treatment of §7.1 of [16] and §3.2 of [15] closely. The subject, initiated by the legendary mathematician Srinivasa Ramanujan, has a fascinating history. We recommend [20] for a short account of it.

Let $w \in \frac{1}{2}\mathbb{Z}$ and let h be a holomorphic function on $\mathbb{H}$ with at most exponential growth at all cusps. We say that h is a (weakly holomorphic) mock modular form of weight w for a discrete subgroup $\Gamma \leq SL_2(\mathbb{R})$ if there is a modular form g of weight $2-w$ such that the sum $\hat{h} := h + g^*$ transforms like a holomorphic modular form of weigh w for Γ. Moreover, we say that g is the *shadow* of the mock modular form h and $\hat{h}$ is its *completion*. In the above we have used the following definition of g^*. Writing the Fourier expansion of g as $g(\tau) = \sum_{n\geq 0} c_g(n)q^n$, then

$$g^*(\tau) := \overline{c_g(0)}\frac{(-\Im(\tau))^{1-w}}{w-1} + \sum_{n>0}(-4\pi n)^{w-1}\overline{c_g(n)}q^{-n}\Gamma(1-w, 4\pi n\Im(\tau)) \ , \tag{2.34}$$

where $\Gamma(1-w,x) = \int_x^\infty e^{-t}t^w dt$ denotes the incomplete gamma function. When $c_g(0) = 0$, the above coincides with the so-called *non-holomorphic weight w Eichler integral* of g, given by

$$g^*(\tau) := (-2)^{w-1}\,\mathrm{e}(\tfrac{w-1}{4})\int_{-\bar{\tau}}^{\infty}(\tau' + \tau)^{-w}\overline{g(-\overline{\tau'})}\mathrm{d}\tau' \ . \tag{2.35}$$

Note that

$$-2i\Im(\tau)^w\frac{\partial}{\partial\bar{\tau}}g^*(\tau) = \overline{g(\tau)} \ , \tag{2.36}$$

and hence $\hat{h}$ is annihilated by the weight w Laplacian $\Delta_w := \Im(\tau)^{2-w}\partial_\tau\Im(\tau)^w\partial_{\bar{\tau}}$. Such functions are called harmonic Maass forms, and one can identify h as the (uniquely defined) holomorphic part of the harmonic Maass form $\hat{h}$. Finally, note that from (2.34) it is obvious that the harmonic Maass form $\hat{h}$ transforms with a multiplier which is the inverse of that of the modular form g.

Just as in the case of usual modular forms, one can generalise the above definition of mock modular forms in various directions, including incorporating non-trivial multiplier systems and considering vector-valued mock modular forms. Next we turn our attention to a specific type of vector-valued mock modular forms, namely those arising from the so-called mock Jacobi forms. For integers k and m, we say that an elliptic form $\phi \in \mathcal{E}_m$ is a *weak mock Jacobi form* of weight k and index m if the following is true. Write the theta-decomposition of ϕ as $\phi = \sum_r h_r\theta_{m,r}$. First, $\tau \mapsto \phi(\tau, z)$ is bounded as $\Im(\tau) \to \infty$ for every fixed $z \in \mathbb{C}$; second, all the h_r are holomorphic; finally, there exists a skew-holomorphic Jacobi form $\sigma = \sum_r \overline{g_r}\theta_{m,r} \in S^{\mathrm{sk}}_{3-k,m}$, such that $\hat{\phi} := \sum_r \hat{h}_r\theta_{m,r}$ is invariant for the weight k modular action $|_{k,m}$ of $SL_2(\mathbb{Z})$ on $\mathcal{E}_m$ (cf. (2.25)) with the definition

$$\hat{h}_r(\tau) := h_r(\tau) + \frac{1}{\sqrt{2m}}g^*_r(\tau) \ . \tag{2.37}$$

As discussed in [21] and analysed carefully in [16, 22], meromorphic Jacobi forms – what one obtains when relaxing the condition on Jacobi forms to allow for poles at torsion points $z \in \mathbb{Q} + \mathbb{Q}\tau$ – naturally give rise to mock Jacobi forms. In particular, all the mock Jacobi forms featured in umbral moonshine can be viewed as arising from meromorphic Jacobi forms.

From a physical point of view, as demonstrated in a series of recent works, the "mockness" of these mock modular objects is often related to the non-compactness of relevant spaces in the theory. See, for instance, [16,23–26]. Let us take 2d CFTs with a non-compact target space as an example. The non-compactness of the target space often lead to a continuous part of the spectrum. In this case the standard CFT arguments might fail. In particular there could be imperfect pairing between the bosonic and fermionic states in the continuous part of the spectrum and we could end up with a non-holomorphic BPS index, given by the completion of a mock modular object, as a result. See for instance [24,27–31] for details for some specific examples, and see the remark at the end of §3.3 for a more detailed discussion in the context of elliptic genus. Another context in which non-compactness appears and leads to a role for mock modular forms is wall-crossing (when approaching the wall, the distance of the bound black hole centers goes to infinity). The BPS counting of the black hole microstates hence depends on the moduli and correspondingly the countour of integration [32], and the result of the integration is mock modular [16].

Another source of mock modular forms in physics is the characters of supersymmetric infinite algebras, such as the $\mathcal{N}=2$ and $\mathcal{N}=4$ superconformal algebras mentioned in §8.1. Some more examples can be found in for instance [33] and references therein. Interestingly, as we will explain in §8.2, the mockness of the mock modular form in (2.42) can be seen as either arising from CFT with non-compact target space or as a result to the mockness of characters of the $\mathcal{N}=4$ superconformal algebra.

We will end this subsection with a few examples.

- Ramanujan wrote down the following simple-looking Eulerian series in his 1920 letter to Hardy [34],

$$\chi_0(q)=\sum_{n=0}^{\infty}\frac{q^n}{(1-q^{n+1})(1-q^{n+2})\dots(1-q^{2n})}=1+q+q^2+2\,q^3+\dots\ ,$$
$$\chi_1(q)=\sum_{n=0}^{\infty}\frac{q^n}{(1-q^{n+1})(1-q^{n+2})\dots(1-q^{2n})(1-q^{2n+1})}=1+2\,q+2\,q^2+3\,q^3+\dots\ ,$$

as two of the examples of his *mock theta functions* (of order 5). In fact, they are closely related to mock Jacobi forms.

Define $I^{3E_8}=\{1,7\}$, $A=\{1,11,19,29\}$, and

$$H_1^{3E_8}(\tau)=q^{-1/120}(2\chi_0(q)-4)\ ,\quad H_7^{3E_8}(\tau)=2q^{71/120}\chi_1(q), \tag{2.38}$$

then $(H_r^{3E_8})_{r\in I^{3E_8}}$ is a vector-valued mock modular form of weight 1/2 for the modular group. Its shadow is given by the index 30 theta functions

$$g_r^{3E_8}=3\sum_{a\in A}\theta^1_{30,a\,r}. \tag{2.39}$$

Writing $\hat{H}_r^{3E_8}(\tau)=H_r^{3E_8}+(g_r^{3E_8})^*$, we have

$$\begin{pmatrix}\hat{H}_1^{3E_8}\\ \hat{H}_7^{3E_8}\end{pmatrix}(\tau+1)=\begin{pmatrix}\mathrm{e}(-\frac{1}{120}) & 0\\ 0 & \mathrm{e}(-\frac{49}{120})\end{pmatrix}\begin{pmatrix}\hat{H}_1^{3E_8}\\ \hat{H}_7^{3E_8}\end{pmatrix}(\tau) \tag{2.40}$$

and

$$\begin{pmatrix} \hat{H}^{3E_8}_1 \\ \hat{H}^{3E_8}_7 \end{pmatrix} \left(-\frac{1}{\tau}\right) = \tau^{1/2} i^{3/2} \begin{pmatrix} \frac{1}{2}\sqrt{1-\frac{1}{\sqrt{5}}} & \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{5}}} \\ \frac{1}{2}\sqrt{1-\frac{1}{\sqrt{5}}} & -\frac{1}{2}\sqrt{1-\frac{1}{\sqrt{5}}} \end{pmatrix} \begin{pmatrix} \hat{H}^{3E_8}_1 \\ \hat{H}^{3E_8}_7 \end{pmatrix} (\tau) \quad (2.41)$$

Moreover, $H^{3E_8}_i$ can be viewed as arising from the theta composition of the mock Jacobi form $\psi^{3E_8}(\tau,z) := \sum_{r\in I^{3E_8}} H^{3E_8}_r \sum_{a\in A} \hat{\theta}^{3E_8}_{a\,r}$. More specifically, $\hat{\psi}^{3E_8}(\tau,z) := \sum_{r\in I^{3E_8}} \hat{H}^{3E_8}_r \sum_{a\in A} \hat{\theta}^{3E_8}_{a\,r}$ is non-holomorphic in τ and transforms as a Jacobi form of weight 1 and index 30. As the notation suggests, $H^{3E_8}_r$ encodes the graded dimension of the umbral moonshine module underlying the case of umbral moonshine corresponding to Niemeier lattice N with root system $3E_8$, as we will discuss in §5.2.

- Let $H : \mathbb{H} \to \mathbb{C}$ be given by

$$H(\tau) = \frac{-2E_2(\tau) + 48F^{(2)}_2(\tau)}{\eta(\tau)^3} = 2q^{-\frac{1}{8}}\left(-1 + 45\,q + 231\,q^2 + 770\,q^3 \ldots\right) \;, \quad (2.42)$$

where

$$E_2 = 1 - 24\sum_{n=1}^{\infty} \frac{nq^n}{1-q^n}$$

is the weight two Eisenstein series (which is not a modular form) and

$$F^{(2)}_2(\tau) = \sum_{\substack{r>s>0 \\ r-s=1 \bmod 2}} (-1)^r\, s\, q^{rs/2} = q + q^2 - q^3 + q^4 + \ldots \;.$$

Note that the first few Fourier coefficients of $H/2$: 45, 231 770, 2277 , 5796, coincide with dimensions of certain irreducible representations of the sporadic group M_{24}! Indeed, in umbral moonshine $H = H^{24A_1}$ plays the role of the graded dimensions of the underlying M_{24}-module. See 5.1.

This function is a mock modular form with shadow $24\eta^3(\tau)$ (and therefore with a multiplier given by the inverse of that of $\eta^3(\tau)$). In other words,

$$\hat{H}(\tau) = H(\tau) + 24\,(4i)^{-1/2}\int_{-\bar{\tau}}^{\infty} (z+\tau)^{-1/2}\overline{\eta(-\bar{z})^3}\mathrm{d}z, \quad (2.43)$$

transforms as a weight 1/2 modular form for the modular group $SL_2(\mathbb{Z})$.

Moreover, the two-variable function $\psi^{24A_1}(\tau,z) := H(\tau)(\theta_{2,1} - \theta_{2,-1})(\tau,z)$ is a mock Jacobi form of weight one and index two. This mock Jacobi form can be seen as arising from a meromorphic Jacobi form by subtracting its "polar part". To see this, consider the weight one index two meromorphic Jacobi form

$$\psi(\tau,z) := -2i\frac{\theta_1(\tau,2z)\eta^3(\tau)}{\theta_1^2(\tau,z)}\phi_{0,1}(\tau,z) = -i\frac{\theta_1(\tau,2z)\eta^3(\tau)}{\theta_1^2(\tau,z)}\mathbf{EG}(\tau,z;K3) \;, \quad (2.44)$$

(cf. (2.31)) which has a simple pole at $z \in \mathbb{Z}+\mathbb{Z}\tau$. Then the following identity holds,

$$\psi(\tau,z) = \psi^{24A_1}(\tau,z) - 24\,\mathrm{Av}^{(2)}\left[\frac{y+1}{y-1}\right] \;. \quad (2.45)$$

In the above $\mathrm{Av}^{(m)}$ denotes the index-m averaging operator

$$\mathrm{Av}^{(m)}[F(y)] = \sum_{k\in\mathbb{Z}} q^{mk^2} y^{2mk} F(q^k y) \ ,$$

with the elliptic symmetry $\mathrm{Av}^{(m)}[F(y)]|_m(\lambda,\mu) = \mathrm{Av}^{(m)}[F(y)]$ for all $\lambda, \mu \in \mathbb{Z}$, and the second term in (2.45) can be interpreted as the canonical "polar part" of the meromorphic Jacobi form φ.

- Consider the set of binary quadratic forms with discriminant D,

  $$\mathcal{Q}_D := \{Q[X,Y] = AX^2 + BXY + CY^2 | B^2 - 4AC = D\} \ . \tag{2.46}$$

  This has a natural action of $SL_2(\mathbb{Z})$, and we are interested in elements in $\mathcal{Q}_D$ that are not equivalent under the modular group action. For instance, an interesting number is the *Hurwitz-Kronecker class number*

  $$H(D) := \sum_{Q\in\mathcal{Q}_D \backslash SL_2(\mathbb{Z})} \frac{1}{\omega_Q} \ , \tag{2.47}$$

  where ω_Q denotes the order of the $SL_2(\mathbb{Z})$-subgroup that leaves Q invariant. It takes values in $\{1, 2, 3\}$. Roughly speaking, this number counts the number of inequivalent quadratic forms with discriminant D, and will play an important role in §6 and §9.

  To each $Q \in \mathcal{Q}_D$ there is a unique root α_D (satisfying $Q[1, \alpha_D] = 0$) in the upper-half plane. Clearly, $J(\alpha_Q)$ is independent of which representative of Q one picks in the $SL_2(\mathbb{Z})$-orbit and one can similarly define weighted sums like (2.47), but now involving values of polynomials of J evaluated at α_Q, referred to as the *traces of singular moduli.* The generating functions of such quantities (summing over D with the grading factor q^{cD} for some c) will turn out to have interesting modular properties.

  For instance, the following generating of the Hurwitz-Kronecker class number, extended by $H(0) = -\frac{1}{12}$,

  $$\mathcal{H}(\tau) := \sum_{D\leq 0} H(D) q^{|D|} = -\frac{1}{12} + \frac{1}{3}q^3 + \frac{1}{2}q^4 + q^7 + q^8 + \dots \tag{2.48}$$

  is a mock modular form of weight $\frac{3}{2}$ for $\Gamma_0(4)$, with shadow given by the theta function $\theta(\tau)$ (see §2.1). The mock modularity property of this function is realised very early on in [35]. This form also has the interesting feature that it is bounded at all cusps and hence has slow growth in its Fourier coefficients, corresponding to the subtle growth of the class numbers.

  The physical relevance of this function will be mentioned in §9. This is in fact perhaps the first mock modular form which was found to play a role in physic [36, 37].

3 Conformal field theory in 2 dimensions

This chapter gives a brief summary of some key ingredients of 2-dimensional conformal field theories (CFTs), and is in no way meant as a complete exposition. CFTs are relevant for moonshine, since in the cases that are known so far the corresponding moonshine modules feature vertex operator algebra (VOA) structures, which capture the structure of the chiral algebra of a 2d CFT. Instead of the more formal VOA language, we opt for the CFT language more familiar to physicists. References on the basic knowledge of CFT include [38–41]. See also Professor Xi Yin's TASI lecture notes in this volume [42].

After the general summary of the basic structure in §3.1, we quickly review aspects of (holomorphic) orbifolds that are relevant for moonshine, in particular for the understanding of the modular properties of the moonshine functions. After that we focus on supersymmetric conformal field theories and introduce the so-called elliptic genus, counting BPS states, that will play an important role in the discussion of the physical context of Mathieu and umbral moonshine (§8).

3.1 General structure

A conformal field theory is a quantum field theory with conformal symmetry. Conformal transformations are coordinate transformations that preserve the conformal flatness of the metric. Focusing on Riemannian manifolds of Euclidean signature, a metric is said to be conformally flat if it can be written in the form $ds^2 = e^{\omega(x)}\delta_{\mu\nu}dx^\mu dx^\nu$. Conformal transformations locally preserve the angles but may deform the lengths arbitrarily, so conformal symmetry is typically associated with the absence of an intrinsic length scale. On the conformal compactification (by adding the point at infinity which is necessary for the special conformal transformation to be well-defined) of $\mathbb{R}^{n,0}$ for $n \geq 3$, all conformal transformations are globally well-defined and form a group isomorphic to $SO(n+1,1,\mathbb{R})$. The corresponding local transformations thus form a finite-dimensional Lie algebra isomorphic to $\mathfrak{so}(n+1,1,\mathbb{R})$. In two dimensions, however, the condition of conformal invariance is equivalent to the Cauchy-Riemann equation and any holomorphic function gives rise to an infinitesimal conformal transformation. Use the generators

$$\ell_n = -z^{n+1}\partial_z \ , \quad \bar{\ell}_n = -\bar{z}^{n+1}\partial_{\bar{z}} \tag{3.1}$$

for $n \in \mathbb{Z}$, we see that the local conformal transformations form an infinite-dimensional Lie algebra, which contains two commuting copies of the Witt algebra with commutation relations

$$\mathfrak{Witt}: \quad [\ell_m, \ell_n] = (m-n)\ell_{m+n} \ . \tag{3.2}$$

It is important to emphasise that most of the conformal generators in 2 dimensions are purely local, i.e. they do not generate globally well-defined transformations. Consider, for example, the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \infty$, i.e. the Riemann surface of genus zero. On $\widehat{\mathbb{C}}$, only $\ell_0, \ell_{\pm 1}$ generate global conformal transformations, which form the Möbius group $SO(3,1,\mathbb{R}) \cong PSL(2,\mathbb{C})$.

The quantisation of a 2-dimensional CFT is typically done on $\mathbb{C}$. The theory on the Riemann sphere determines the theory on any other Riemann surface uniquely, but does not guarantee their consistency, as one must also require crossing symmetry and modular invariance of the torus partition function (see below). To see how to quantise on $\mathbb{C}$, we note that $\mathbb{C}$ with origin removed is conformally equivalent to a cylinder $S^1 \times \mathbb{R}$. Denoting by y and t the coordinates for S^1 and the Euclidean time $\mathbb{R}$, the conformal map $z = e^{t+iy}$ maps the cylinder to $\mathbb{C}\backslash\{0\}$ and in particular maps the infinite past to the origin. The usual time ordering on the cylinder becomes radial ordering on the plane, and the associated space of states is built on radial slices.

Anything that resembles a local field $\phi(z, \bar{z})$ is called a field in CFT. If a field depends only on the holomorphic variable z we call it *chiral field* (or *anti-chiral* if it depends only on $\bar{z}$). Upon quantisation, fields become operator-valued distributions that create states in the space of states $\mathcal{H}$, by acting on the vacuum $|0\rangle \in \mathcal{H}$. This is called the *state-field correspondence*, which maps an field ϕ to a state

$$\phi \mapsto |\phi\rangle := \lim_{z,\bar{z}\to 0} \phi(z, \bar{z})|0\rangle \ , \tag{3.3}$$

created at the origin on the plane (or past infinity on the cylinder). A crucial property of a CFT is that the above map is bijective; every state corresponds uniquely to a single local operator, whereas for a typical QFT different fields can produce the same asymptotic state. This can be understood through the fact that under conformal transformation $t \to -\infty$ is mapped to a single local point on $\widehat{\mathbb{C}}$.

The product of two fields inserted at the same point is generically singular. The singularity structure is captured by the so-called *operator product expansion* (OPE)

$$\phi_1(z)\phi_2(z') \sim \sum_{n=0}^{\infty} D_n(z - z')O_n(z') \ , \tag{3.4}$$

where $\sim$ means that we only keep the singular terms. Here $O_n(z)$ are fields of the theory and $D_n(z - z')$ are complex-valued functions with polynomial or logarithmic singularities when $z \to z'$. The non-singular part of $\phi_1(z)\phi_2(z')$ is captured by the *normal-ordered product*, which can be defined as

$$:\!\phi_1(z)\phi_2(z')\!: \ := \phi_1(z)\phi_2(z') - \sum_{n=0}^{\infty} D_n(z - z')O_n(z') \ . \tag{3.5}$$

When there are only polynomial singularities in $D_n(z - z')$ we say that the fields ϕ_1 and ϕ_2 are local with respect to each other, in the sense that there are no branch cuts and contour integrals are well-defined.

The conserved current associated with the continuous conformal symmetry of a 2d CFT is the *stress-energy tensor*, and we denote $T(z) := T_{zz}(z)$ and $\bar{T}(\bar{z}) := T_{\bar{z}\bar{z}}(\bar{z})$. Classically, these are the only non-vanishing components. Upon quantisation on a generic Riemann surface this is broken to $\langle T^a_{\ a}\rangle = -\frac{c}{12}R$ where R is the Ricci scalar.

Since the treatment of the chiral and anti-chiral parts is identical, we will from now on focus on the former. The holomorphicity of the stress-energy tensor $T(z)$ follows from the fact that the associated conserved charges are precisely the generators of infinitesimal holomorphic conformal transformations (3.1). Specifically, we have the mode expansion

$$L_n := \frac{1}{2\pi} \oint T(z) z^{n+1} dz \ \Leftrightarrow \ T(z) = \sum_{n\in\mathbb{Z}} L_n z^{-n-2} \ . \tag{3.6}$$

The modes L_n however, do not generally satisfy the Witt algebra (3.2). This is because the conformal symmetry is typically "softly" broken by quantum effects. The OPE of $T(z)$ with itself,

$$T(z)T(w) \sim \frac{c/2}{(z - w)^4} + \frac{2T(w)}{(z - w)^2} + \frac{\partial T(w)}{z - w} \ , \tag{3.7}$$

is equivalent via the mode expansions to the commutation relations

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n,0} \ . \tag{3.8}$$

In the above, the real constant c is called *central charge*, and the new algebra is the *Virasoro algebra* $\mathfrak{Vir}$, which is a central extension of $\mathfrak{Witt}$ by the term containing the central charge. Moreover, the two resulting $\mathfrak{Vir}$ copies commute, i.e. $[L_m, \bar{L}_n] = 0$, and there is a central charge $\bar{c}$ associated with the anti-chiral part, which can in principle be different from c. The central charge captures important information of a CFT and gives a measure for the number of degrees of freedom, but there can exist multiple different CFTs with the same central charge. It is related to a "soft" breaking of the conformal symmetry because it indicates that the stress-energy tensor, which generates conformal transformations, transforms anomalously under conformal mappings. For instance, for the transformation $z = e^w, w = t + iy$ from the cylinder to the Riemann sphere, we have

$$T(z) = z^{-2}\left(T_{\text{cyl}}(w) + \frac{c}{24}\right) \ , \tag{3.9}$$

which leads to the following relation between the Hamiltonian $\oint_{\text{time-slice}} dw (T_{\text{cyl}}(w) + \bar{T}_{\text{cyl}}(\bar{w}))$ and $\oint_{\text{radial-slice}} \frac{dz}{z}(T(z) + \bar{T}(\bar{z}))$:

$$H = L_0 + \tilde{L}_0 - \frac{c+\tilde{c}}{24}. \tag{3.10}$$

Similarly, we have the momentum (or *spin*)

$$P = L_0 - \tilde{L}_0 - \frac{c-\tilde{c}}{24}. \tag{3.11}$$

As a result, the eigenvalues of L_0 plays the role of the chiral part of the energy, and the central charge gives rise to non-vanishing ground state energy. The eigenvalue h under L_0 of an eigenstate $|h\rangle \in \mathcal{H}$, i.e. $L_0|h\rangle = h|h\rangle$, is called the *conformal weight* of $|h\rangle$. If, moreover, $L_n|h\rangle = 0$ for all $n > 0$, then $|h\rangle$ is called a *Virasoro primary state* and the corresponding field called a primary field. This terminology also extends to the corresponding fields ϕ_h via the state-field correspondence. A state of the form $L_{-k_1}L_{-k_2}\cdots L_{-k_n}|h\rangle$ $(k_i > 0)$ is called a *Virasoro descendant* of $|h\rangle$. If $|h\rangle$ is a primary state, then along with all of its descendants they form a so-called *Verma module* for $\mathfrak{Vir}$. The primary state $|h\rangle$ is then called the *highest-weight state* of the module, since it has the lowest (somewhat confusingly) conformal weight among all of its descendants.

Since the states organise themselves into Virasoro representations, one can decompose the space of states of a CFT into a direct sum of $\mathfrak{Vir}$ and $\overline{\mathfrak{Vir}}$ modules. In general, focussing on the chiral part, one can have an enlarged symmetry algebra that contains $\mathfrak{Vir}$. This is called the *chiral algebra* of the CFT, and is denoted here by $\mathcal{V}$. We are mainly interested in *rational conformal field theories* (RCFTs), which contain a finite number of such modules; let $\Phi(\mathcal{V})$, $\overline{\Phi}(\overline{\mathcal{V}})$ denote the sets of these (chiral and anti-chiral respectively). The space of states can then be written as

$$\mathcal{H} \ = \bigoplus_{M\in\Phi(\mathcal{V}),\overline{N}\in\overline{\Phi}(\overline{\mathcal{V}})} \mathcal{Z}_{M,\overline{N}}\left(M \otimes \overline{N}\right) \ . \tag{3.12}$$

The states in such RCFTs get organized in $\mathcal{V}$-modules, whose highest-weight states correspond to *chiral primaries*, which are not only Virasoro primaries, but also primaries with respect to $\mathcal{V}$. The chiral descendants are generated by acting with $\mathcal{V}$ on the chiral primaries. This means that if $A(z) \in \mathcal{V}$, then $A(z)M \subset M$ for any $M \in \Phi(\mathcal{V})$.

The subspace $V \subset \mathcal{H}$, corresponding to the chiral algebra $\mathcal{V}$ via the state-field correspondence, always forms an irreducible $\mathcal{V}$-module $V := M_1 \in \Phi(\mathcal{V})$, which contains the vacuum and all states corresponding to the Virasoro primaries generating $\mathcal{V}$, also commonly called *currents*, along with their chiral descendants. Modular invariance requires that the eigenvalues of $L_0 - \overline{L}_0$ are integers, which also means that all states in any $M \in \Phi(\mathcal{V})$ have the same weight up to an integer. Specifically, all states in the vacuum module V should have integer weights. However, by dropping modular invariance as an initial requirement, the chiral algebra can possibly contain currents of half-integer weights (fermionic currents), or any rational weight (parafermions). The price to pay is that these currents have non-local OPEs (in the sense discussed previously), with the corresponding branch cuts leading to the introduction of various sectors (for fermionic currents these would be the Ramond and Neveu-Schwarz sectors). The modular invariant theory can then be constructed by a suitable projection.

The chiral algebras themselves are the central objects in the theory of Vertex Operator Algebras (VOAs), where they are discussed in an axiomatic manner. In the context of moonshine, an important property of a chiral algebra is the finite group part of $\mathrm{Aut}(\mathcal{V})$. The most famous example is the Monster CFT $V^\natural$, which is a VOA with $\mathrm{Aut}(V^\natural) = \mathbb{M}$, the Monster group. Furthermore, $V^\natural$ is an example of a *holomorphic VOA*, i.e. a VOA that has a unique irreducible $\mathcal{V}$-module, namely the space V.

In 2d CFT one is interested in calculating correlation functions of fields, inserted at specific points on a Riemann surface Σ. These can be cast in terms of chiral quantities called *chiral blocks*. Writing $\Sigma = \Sigma_{g,n}$ with genus g and n marked points $p_1, \dots, p_n$, a chiral block is a multilinear map from $M_1 \otimes \cdots \otimes M_n$ to a meromorphic function. This notation means that a field in $M_i \in \Phi(\mathcal{V})$ is inserted at the point p_i. In the case of RCFT, they can often be obtained as solutions to certain differential equations [43–45]. The chiral blocks form representations of the *mapping class group* $\Gamma_{g,n}$, which captures the discrete (and almost always infinite) symmetries of $\Sigma_{g,n}$. It can be defined by the quotient $\Gamma_{g,n} \cong \mathrm{Aut}(\Sigma_{g,n})/\mathrm{Aut}_0(\Sigma_{g,n})$, where $\mathrm{Aut}_0(\Sigma_{g,n})$ is the component of $\mathrm{Aut}(\Sigma_{g,n})$ that is connected to the identity. Hence, $\Gamma_{g,n}$ maps between equivalent Riemann surfaces $\Sigma_{g,n}$, which only differ by a discrete automorphism. As a result, the *moduli space* $\mathcal{M}_{g,n}$, which parametrises the conformally inequivalent Riemann surfaces, has naturally the following quotient form,

$$\mathcal{M}_{g,n} = \mathcal{T}_{g,n}/\Gamma_{g,n} \ , \tag{3.13}$$

where $\mathcal{T}_{g,n}$ is the so-called *Teichmüller space*.

Chiral blocks have in general non-trivial monodromy as functions of the moduli space $\mathcal{M}_{g,n}$ (see for example [40] for more details). Chiral blocks will thus generally be multi-valued functions on $\mathcal{M}_{g,n}$, and in order to make them well-defined one should lift them to $\mathcal{T}_{g,n}$. As a result, they will then carry a representation of the mapping class group $\Gamma_{g,n}$. This is one way to understand the origin of the modular properties of torus blocks, and in particular moonshine modules.

To explain this, let us now focus on the case of $\Sigma_{1,1}$, i.e. tori with a single marked point. As explained in §2.1, a torus can be described up to a scale by $\mathbb{C}/\Lambda_\tau$, where Λ_τ is the lattice in $\mathbb{C}$ generated by the vectors 1 and $\tau \in \mathbb{H}$. An

$SL_2(\mathbb{Z})$ transformation leaves the lattice invariant and as a result the mapping class group $\Gamma_{1,0} = PSL_2(\mathbb{Z})$ is given by the part of $SL_2(\mathbb{Z})$ that acts non-trivially on the Teichmüller space $\mathbb{H}$. As any point is equivalent to any other point on a torus due to its translation symmetries, we also have $\Gamma_{1,0} = \Gamma_{1,1}$ and $\mathcal{M}_{1,0} = \mathcal{M}_{1,1}$.

Chiral blocks on $\Sigma_{1,1}$, when lifted to $\mathcal{T}_{1,1}$, will consequently be functions of the modular parameter τ. For RCFTs, they form a space of finite dimensions, and the dimension is given by the number of irreducible modules in $\Phi(\mathcal{V})$. They admit a natural basis given by the *graded dimensions*, or *characters*, of the irreducible modules $M \in \Phi(\mathcal{V})$

$$ch_M(\tau) = \mathrm{Tr}_M \, q^{L_0 - c/24} \ , \tag{3.14}$$

where $q = e^{2\pi i \tau}$ as before. As discussed previously, the characters furnish a representation of $\Gamma_{1,1} = PSL(2,\mathbb{Z})$, so that the $ch_M(\tau)$ are components of a weakly holomorphic vector-valued modular form for $PSL(2,\mathbb{Z})$. In other words, they mix with each other under the action of the modular group and the way they mix determines their OPE via the Verlinde formula. The modularity of characters of RCFTs is rigorously shown in the context of VOAs by Zhu's Theorem [46].

The *partition function* of a 2d CFT is defined as the 0-point correlation function on the torus, which encodes the spectrum of the theory. In the operator formalism, a torus with modular parameter $\tau = \tau_1 + i\tau_2$ can be obtained from the Riemann sphere by first conformally mapping it to the cylinder $S^1 \times \mathbb{R}$, and then imposing periodic boundary conditions on the Euclidean time direction $\mathbb{R}$. The Hamiltonian and momentum operators H, P then propagate states along both cycles of the torus, so the spectrum is embodied in the trace of the corresponding evolution operator over the space of states,

$$Z(\tau,\bar\tau) := \mathrm{Tr}_{\mathcal{H}} \, e^{2\pi i \tau_1 P - 2\pi \tau_2 H} \ . \tag{3.15}$$

Using (3.10)-(3.11), we can rewrite it as

$$Z(\tau,\bar\tau) = \mathrm{Tr}_{\mathcal{H}} \, q^{L_0 - \frac{c}{24}} \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} \ , \tag{3.16}$$

making it manifest that it is a generating function of the multiplicities of states at given chiral and anti-chiral conformal weights in $\mathcal{H}$. From (3.12) we see that it has the following decomposition in terms of chiral blocks

$$Z(\tau,\bar\tau) = \sum_{M \in \Phi(\mathcal{V}), \overline{N} \in \overline{\Phi}(\overline{\mathcal{V}})} \mathcal{Z}_{M,\overline{N}} \ ch_M(\tau) \overline{ch_{\bar{N}}(\tau)} \ . \tag{3.17}$$

The partition function (3.15) can also be computed using the path integral formalism when a Lagrangian description of the CFT is available. In this language, we have $Z = \int D\phi \ e^{-S[\phi]}$, with the fields having appropriate boundary conditions on the two cycles of the torus. Also from this point of view, it is clear that the partition function should be modular invariant. This invariance imposes severe constraints on the spectrum of 2d CFTs. For instance, modular invariance was used to classify supersymmetric minimal models and further extensions. See [47, 48] and references therein for some of these results. In the context of moonshine, we are mainly interested in the chiral CFT, where the modular properties are not as stringent.

3.2 Orbifolds

A special class of CFTs which is of particular interest for moonshine is the so-called *orbifold CFTs* [49]. The orbifold construction essentially entail "gauging" a discrete symmetry group G of the chiral algebra $\mathcal{V}$. More precisely, it builds a theory whose chiral algebra contains the G-invariant subalgebra $\mathcal{V}^G$ of $\mathcal{V}$, by retaining the G-invariant states of the original theory and introducing new "g-twisted" sector states, for every $g \in G$.

There are two important ways orbifold considerations enter the study of moonshine. First, we will see in §4 explicit constructions of moonshine chiral CFTs obtained by $\mathbb{Z}_2$-orbifolds. Second, the partition functions twined by the finite group symmetries provide the necessary information about the group actions on the moonshine CFT and constitute the modular objects playing a central role in moonshine. Generalising this to the twisted sectors leads to the so-called generalised moonshine, which we will mention in the next part of the lecture.

Orbifold chiral algebra. Here we are mainly interested in orbifolds of chiral RCFTs (rational VOAs). We are interested in automorphisms of the operator algebra. If such an automorphism acts trivially on the operator algebra, i.e. without permuting the modules M_i^e, then it is said to be *inner*. In particular it preserves the chiral algebra of the chiral CFT. Let $\mathcal{V}$ denote the chiral algebra, $G \subseteq \mathrm{Aut}(\mathcal{V})$ a finite symmetry group, and $M_1^e, \ldots, M_n^e$ its irreducible $\mathcal{V}$-modules. Here $e \in G$ denotes the identity element which will later be generalised to arbitrary $g \in G$. In particular, we have $M_1^e = V$, the vacuum module corresponding to $\mathcal{V}$.

Given such a symmetry, the chiral algebra is decomposed in G-representations as

$$\mathcal{V} = \bigoplus_a \rho_a \otimes \mathcal{V}_a \ , \tag{3.18}$$

where the corresponding spaces V_a contain states that transform under the irreducible representations ρ_a of G, and a runs over all of them. The G-invariant subalgebra

$$\mathcal{V}^G := \{\phi \in \mathcal{V} \mid h\phi = \phi \ \ \forall\, h \in G\} \ , \tag{3.19}$$

corresponding to the trivial representation ρ_0 of G, is called the *orbifold chiral algebra* in this setup. Note that while V is irreducible as a $\mathcal{V}$-module, it is generically reducible as a $\mathcal{V}^G$-module, as shown in (3.18). We instead identify the corresponding space V_a, corresponding to $\mathcal{V}_a$, as the irreducible $\mathcal{V}^G$-modules relevant for the orbifold CFT.

An analogous statement holds for the rest of the $\mathcal{V}$-modules, and we have decompositions

$$M_i^e = \bigoplus_a \rho_a \otimes M_{i,a}^e \ . \tag{3.20}$$

An important subtlety is that ρ now runs over all irreducible *projective representations* of G. Projective representations generalise the usual notion of representations introduced in §1.4, by allowing them to respect the group operation up to a phase,

$$\rho(h_1 h_2) = c_e(h_1, h_2)\rho(h_1)\rho(h_2) \ , \tag{3.21}$$

where $c_e(h_1, h_2)$ is a $U(1)$-valued 2-cocycle, representing a class in the group cohomology $H^2(G, U(1))$ of G. This type of behaviour is allowed in CFT because such a phase cancels when the chiral and the anti-chiral contributions are combined and hence is not in conflict with the modular invariance of the final theory. See [50] for a nice survey on projective representations of finite groups. Also note that the G-invariance of the vacuum implies that the vacuum module V carries true representations in the decomposition $V = \bigoplus_a \rho_a \otimes V_a$.

Twinings. For each $h \in G$, acting as an inner automorphism of the operator algebra, we define the *twined characters*

$$ch_i \left(h \underset{e}{\square} ; \tau \right) := \mathrm{Tr}_{M_i^e} \left[h \; q^{L_0 - \frac{c}{24}} \right] \ . \tag{3.22}$$

Note that the the special case $h = e$ simply gives the usual character or *graded-dimensions*, of M_i^e. In terms of the decomposition into irreducible $\mathcal{V}^G$-modules, the twined characters are expressed as

$$ch_i \left(h \underset{e}{\square} ; \tau \right) = \sum_a \chi_a(h) \; ch \left(M_{i,a}^e \right) \ , \tag{3.23}$$

where χ_a are projective characters of G, and $ch \left(M_{i,a}^e \right)$ are the graded dimensions of $M_{i,a}^e$. Using the orthogonality of the projective representations analogous to (1.18) one can obtain the character $ch_i \left(h \underset{e}{\square} ; \tau \right)$ from the 2-cycle and the character of the projective representation ρ_a.

In a similar fashion, we define the twined partition function as

$$Z \left(h \underset{e}{\square} ; \tau, \bar{\tau} \right) := \mathrm{Tr}_{\mathcal{H}} \left[h \; q^{L_0 - \frac{c}{24}} \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} \right] \ . \tag{3.24}$$

In the path integral language, the twined partition function is obtained by imposing h-twisted boundary condition for the fields on the cycle of the torus which is identified with the "temporal" circle, while the boundary condition along the spatial circle remains unchanged, i.e. $\phi(z + \tau) = h \cdot \phi(z)$ and $\phi(z + 1) = \phi(z)$. From this point of view, it is clear that $Z \left(h \underset{e}{\square} ; \tau, \bar{\tau} \right)$ should be invariant under a subgroup of $SL_2(\mathbb{Z})$ that preserves the h-twisted boundary condition ($SL_2(\mathbb{Z})$ transforms the boundary conditions on the two independent cycles of the torus as in (3.27) below).

Twisted sectors. Provided that $\mathcal{V}$ is sufficiently nice, in the sense that it satisfies the so-called C_2-cofiniteness condition (see [40] for the definition), then for any $g \in \mathrm{Inn}(\mathcal{V})$, the inner automorphisms of $\mathcal{V}$, one can define an irreducible *g-twisted $\mathcal{V}$-module* M_i^g for each $i = 1, \cdots, n$ [51]. In an orbifold theory these modules make up the g-twisted sector of the theory. Clearly, G is no longer a symmetry group for these modules; only the centraliser subgroup $C_G(g)$ (cf. §1.1) remains as a symmetry of the g-twisted sector. As a result, for any commuting pair $g, h \in G$, we can analogously define the *twisted-twined characters*

$$ch_i \left(h \underset{g}{\square} ; \tau \right) := \mathrm{Tr}_{M_i^g} \left[h \; q^{L_0 - \frac{c}{24}} \right] \ , \tag{3.25}$$

i.e. the twined characters in the twisted sectors (of which (3.22) is a special case). In the path integral language, they are obtained by additionally imposing g-boundary conditions for the spacial cycle of the torus, i.e. we have $\phi(z + 1) = g \cdot \phi(z)$ as well as $\phi(z + \tau) = h \cdot \phi(z)$. They similarly admit the decomposition

$$M_i^g = \bigoplus_a \; \rho_a \otimes M_{i,a}^g \ , \tag{3.26}$$

where the sum now runs over all irreducible projective representations of $C_G(g)$. Accordingly, an obvious generalisation of (3.23), obtained by replacing e with g and G with $C_G(g)$, also holds for the twisted sectors.

We have already mentioned that the twined partition functions enjoy modular properties. Similarly the characters (3.25) form vector-valued modular forms for some congruence subgroup with certain multiplier systems. This can be understood via the $SL_2(\mathbb{Z})$-action on the boundary conditions: under modular transformations $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$ on the torus, the boundary conditions (g,h) on the two cycles change as

$$(g,h) \mapsto \left(h^c g^d, h^a g^b\right) \ . \tag{3.27}$$

As a result, the twisted-twined characters transform as

$$ch_i\left(h\underset{g}{\square}\,;\frac{a\tau+b}{c\tau+d}\right) = \sum_{j=1}^{n} \psi(\gamma,g,h)_{ij}\ ch_i\left(h^a g^b\underset{h^c g^d}{\square}\,;\tau\right) \ , \tag{3.28}$$

where $\psi(\gamma,g,h)$ is an $n \times n$ matrix with scalar entries. The special case of *holomorphic VOAs* , i.e. those that contain only a single irreducible (untwisted) $\mathcal{V}$-module $\mathcal{H}$, is the easiest to describe. In this case, the chiral partition function coincides with the character of the chiral algebra and (3.28) becomes [52]

$$\begin{aligned} Z\left(h\underset{g}{\square}\,;\tau+1\right) &= c_g(g,h) Z\left(gh\underset{g}{\square}\,;\tau\right) \ , \\ Z\left(h\underset{g}{\square}\,;-\frac{1}{\tau}\right) &= \overline{c_h\,(g,g^{-1})} Z\left(g^{-1}\underset{h}{\square}\,;\tau\right) \ , \end{aligned} \tag{3.29}$$

where now the phases are given by a 2-cocycle representing a class in $H^2(C_G(g),U(1))$ as in (3.21). Moreover, all the phases for all g should descend from a 3-cocycle representing a class in $H^3(G,U(1))$ [52,53].

In a non-chiral CFT, the spectrum consists of the G-invariant parts of all the twisted sectors, leading to the following expression for the partition function

$$Z(\tau,\overline{\tau}) = \frac{1}{|G|} \sum_{gh=hg} Z(g,h)\epsilon(g,h) \tag{3.30}$$

where $\epsilon(g,h)$ is a phase called the *discrete torsion*, which is just 1 in the simplest cases of orbifold constructions. As usual, the above partition function is modular invariant in a consistent orbifold CFT.

3.3 Elliptic genus

In the previous subsections we have discussed conformal theories in general. In the context of string theory and in this lecture, we often encounter 2d CFTs with supersymmetries. In this subsection, we will start with introducing some necessary background on superconformal algebras and their representations, and in particular explain what an elliptic genus is, first from a physics point of view and then from a geometric point of view.

With supersymmetries, the presence of fermions leads to many new features, stemming from the fact that there is now an extra $\mathbb{Z}_2$ grading on the Hilbert space: $V = V_0 \oplus V_1$. (In the context of moonshine, this leads to supermodules of finite groups, cf. (1.20).) For instance, in the context of type II superstrings compactified on Calabi-Yau manifolds, the relevant "internal" CFT is a non-linear sigma model with $\mathcal{N} = 2$ supersymmetry. The Calabi-Yau structure of the target space guarantees that the theory has the $\mathcal{N} = 2$ extension of Virasoro symmetry, given by the so-called $\mathcal{N} = 2$ superconformal algebra (SCA). In particular, superstrings on $K3$ manifolds and the corresponding elliptic genus will play an important role in §8.

The terminology "$\mathcal{N} = 2$" refers to the fact that we include 2 fermionic currents in the algebra on top of the bosonic energy-momentum tensor $T(z)$. Furthermore, there's now an extra automorphism, called the R-symmetry, that rotates different fermionic currents onto each other.

We denote the two fermionic currents by $G_+(z)$ and $G_-(z)$ and the $U(1)$ R-symmetry current rotating the two by $J(z)$. The algebra reads

$$\begin{aligned}
[L_m, L_n] &= (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\,\delta_{m+n,0} \\
[J_m, J_n] &= \frac{c}{3}m\,\delta_{m+n,0} \\
[L_n, J_m] &= -m\,J_{m+n} \\
[L_n, G_r^\pm] &= (\frac{n}{2}-r)\,G_{r+n}^\pm \\
[J_n, G_r^\pm] &= \pm G_{r+n}^\pm \\
\{G_r^+, G_s^-\} &= 2L_{r+s} + (r-s)J_{r+s} + \frac{c}{3}(r^2 - \frac{1}{4})\,\delta_{r+s,0}\ ,
\end{aligned} \tag{3.31}$$

and all other (anti-)commutators are zero. As before we have two possible boundary conditions for the fermions

$$\begin{cases} 2r = 0 \bmod 2 & \text{for R sector} \\ \\ 2r = 1 \bmod 2 & \text{for NS sector}\ . \end{cases} \tag{3.32}$$

Two comments about this algebra are in order here. First, we have now two generators, L_0 and J_0, of the Cartan subalgebra. As a result, the representations will now be graded by two "quantum numbers", given by the eigenvalues of the L_0 and J_0 of the highest weight vector. The second new feature is that there is a non-trivial inner automorphism of the algebra, which means that the algebra remains the same under the following redefinition

$$\mathbf{SF}_\eta :\ \begin{aligned} L_n &\mapsto L_n + \eta J_n + m\,\eta^2\,\delta_{n,0} \\ J_n &\mapsto J_n + 2m\,\eta\,\delta_{n,0} \\ G_r^\pm &\mapsto G_{r\pm\eta}^\pm \end{aligned} \tag{3.33}$$

with $\eta \in \mathbb{Z}$. This automorphism is called *spectral flow*, and in the above we have written $m := c/6$. If instead we choose $\eta \in \mathbb{Z} + 1/2$ we exchange the Ramond and the Neveu-Schwarz algebra. Note that the only operator (up to trivial rescaling and the addition of central terms, of course) invariant under such a transformation is $4mL_0 - J_0^2$. Recall also that NS sector states give spacetime fermions and Ramond sector states give spacetime bosons. Hence the spectral flow operator has an intimate relation to spacetime supersymmetries.

Ramond ground states and the Witten index. In what follows we will focus on the Ramond algebra and define the Ramond ground states of $\mathcal{N} = 2$ SCFT. As usual, we require the ground states to be annihilated by all the positive modes:

$$L_n|\phi\rangle = J_m|\phi\rangle = G_r^\pm|\phi\rangle = 0 \quad \text{for all} \quad m, n, r > 0\ .$$

Moreover, they have to annihilated by the zero modes of the fermionic currents

$$G_0^{\pm}|\phi\rangle = 0\ .$$

This condition fixes their L_0-eigenvalue to be

$$\frac{1}{2}\{G_0^+, G_0^-\}|\phi\rangle = \Big(L_0 - \frac{c}{24}\Big)\,|\phi\rangle = 0\ .$$

Let's ignore the right-moving part of the spectrum for a moment and consider a chiral Hilbert space V. We define its Witten index as

$$\mathrm{WI}(\tau, V) = \mathrm{Tr}_V\left((-1)^{J_0} q^{L_0 - \frac{c}{24}}\right)\ .$$

If a state $|\psi\rangle$ is not annihilated by G_0^+, then the states $|\psi\rangle$ and $G_0^+|\psi\rangle$ together contribute 0 to $\mathrm{WI}(\tau, V)$ since $[L_0, G_0^+] = 0$ while $[J_0, G_0^+] = G_0^+$. The same argument holds for G_0^- and we conclude that only Ramond ground states can contribute to the Witten index. As a result, the Witten index WI : $\{\mathcal{N} = 2 \text{ SCFT}\} \to \mathbb{Z}$ is independent of τ and counts (with signs) the number of Ramond ground states in V.

Notice moreover that the Witten index for $\mathcal{N} = 2$ SCFT acquires an interpretation as computing the graded dimension of the cohomology of the G_0^+ operator, satisfying $(G_0^+)^2 = 0$. For $\{G_0^+, (G_0^+)^\dagger\} = \{G_0^+, G_0^-\} = L_0 - \frac{c}{24}$, the Ramond ground states have the interpretation as the harmonic representative in the cohomology. This fact underlies the rigidity property of the Witten index and the elliptic genus which we will define now.

The same analysis can be trivially extended when one has a non-chiral theory with both left- and right-moving degrees of freedom: the Witten index

$$\mathrm{WI}(\tau, \bar\tau, V) = \mathrm{Tr}_V\left((-1)^{\tilde J_0 + J_0} \bar q^{\tilde L_0 - \frac{\tilde c}{24}} q^{L_0 - \frac{c}{24}}\right)$$

counts states that are Ramond ground states for both the left- and the right-moving copy of $\mathcal{N} = 2$ SCA.

The $\mathcal{N} = 2$ elliptic genus. It is fine to be able to compute the graded dimension of a cohomology, but we can go further and compute more interesting properties of this vector space. For instance, we have learned that the representations of $\mathcal{N} = 2$ SCA are labelled by two quantum numbers corresponding to the Cartan generators L_0 and J_0. It will hence be natural to consider the following quantity which computes the dimension of $\tilde G_0^+$ cohomology graded by the left-moving quantum numbers L_0, J_0.

The elliptic genus of a $\mathcal{N} = (2, 2)$ SCFT is defined as the following Hilbert space trace

$$\mathbf{EG}(\tau, z) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{RR}}}\left((-1)^{J_0 + \bar J_0} y^{J_0} q^{L_0 - c/24} \bar q^{\bar L_0 - c/24}\right)\ , \quad y = e^{2\pi i z}\ , \tag{3.34}$$

where $\mathcal{H}_{\mathrm{RR}}$ denotes the Hilbert space of states that are in the Ramond sector of the $\mathcal{N} = 2$ SCA both for the left- and right-moving copy of the algebra. From the same argument as that for the Witten index, this quantity will be independent on $\bar q$ and will hence be holomorphic as a function of both τ and z.

Note that the elliptic genus can be seen as something between the partition function and the Witten index. While the former counts all states and the latter counts only RR ground states, the elliptic genus counts states that are Ramond ground state on one side and unconstrained on the other side. It contains a lot more information but still has the rigidity property of the Witten index which makes it possible to compute for many SCFTs, and as such it offers a good balance between information content and computability.

When the theory has a finite group symmetry G which commutes with the superconformal symmetries, one can define the ellpitic genus twined by $g \in G$ as

$$\mathbf{EG}_g(\tau, z) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{RR}}}\left(g\, (-1)^{J_0+\bar{J}_0} y^{J_0} q^{L_0 - c/24} \bar{q}^{\bar{L}_0 - c/24} \right). \tag{3.35}$$

These objects will play an important role in the discussions in Part II and III.

Modular properties. As in the case of partition functions (cf. §3.1), a path integral interpretation of the elliptic genus suggests it has nice transformation property under the torus mapping class group. Moreover, the inner automorphism of the algebra (the spectral flow symmetry) implies that the graded dimension of a L_0-, J_0- eigenspace should only depends on its eigenvalue under the eigenvalue of the combined operator $4mL_0 - J_0^2$ and the charge of J_0 mod $2m$ where $m = c/6$. Hence, the Fourier expansion of the elliptic genus should take the form

$$\mathbf{EG}(\tau, z) = \sum_{n,\ell} q^n y^\ell \, c(4mn - \ell^2, \ell) \ .$$

where $c(D, \ell)$ only depends on D and ℓ mod $2m$ (cf. (2.26)). From these facts one can deduce that the elliptic genus of an $\mathcal{N} = (2, 2)$ SCFT with central charge $c = 6m$ is a weak Jacobi form of weight zero and index m. Similarly, following the same argument and that in §3.2, the twined elliptic genera are also weak Jacobi form of weight zero and the same index, but with the modular group $SL_2(\mathbb{Z})$ in (2.25) replaced by a certain subgroup which depends on the twining symmetry g.

The geometric elliptic genus. For a compact complex manifold M with $\dim_{\mathcal{C}} M = d_0$, we can define its elliptic genus as the character-valued Euler characteristic of the infinite-dimensional formal vector bundle [54–58]

$$\mathbf{E}_{q,y} = y^{d/2} {\textstyle\bigwedge}_{-y^{-1}} T_M^* \bigotimes_{n\geq 1} {\textstyle\bigwedge}_{-y^{-1}q^n} T_M^* \bigotimes_{n \geq 1} {\textstyle\bigwedge}_{-yq^n} T_M \bigotimes_{n\geq 0} S_{q^n}(T_M \oplus T_M^*),$$

where T_M and T_M^* are the holomorphic tangent bundle and its dual, and we adopt the notation

$${\textstyle\bigwedge}_q V = 1 + qV + q^2 {\textstyle\bigwedge}^2 V + \dots, \quad \text{and} \quad S_q V = 1 + qV + q^2 S^2 V \dots,$$

with $S^k V$ denoting the k-th symmetric power of V. In other words, we have

$$\mathbf{EG}(\tau, z; M) = \int_M ch(\mathbf{E}_{q,y}) \mathrm{Td}(M). \tag{3.36}$$

From the above definition we see that this “stringy” topological quantity reduces to the familiar ones: the Euler number, the signature, and the $\hat{A}$ genus of M, when we specialise z to $z = 0, \tau/2, (\tau + 1)/2$, respectively.

When M has vanishing first Chern class, in particular when M is a Calabi–Yau manifold, its elliptic genus $\mathbf{EG}(\tau, z; M)$ can be shown to be a weak Jacobi form of weight zero and index $d_0/2$ [58]. Note that the supersymmetric sigma model on a Calabi–Yau manifold flows to a superconformal SCFT in the infrared. The elliptic genus of this $\mathcal{N} = (2, 2)$ SCFT, defined as in (3.34), then coincides with the geometric elliptic genus defined in (3.36) of the Calabi–Yau manifold.

Examples: $K3$ and T^4. There are two topologically distinct Calabi-Yau two-folds: $K3$ and T^4. We expect their elliptic genus to be weight zero weak Jacobi forms with index 1. Coincidentally, the space of such a form is one-dimensional and is spanned by $\phi_{0,1}(\tau, z)$ (cf. (2.31)), and hence we only need one topological invariant of the Calabi-Yau two-folds to fix the whole elliptic genus. From

$$\mathbf{EG}(\tau, z = 0; T^4) = \chi(T^4) = 0 \quad , \quad \mathbf{EG}(\tau, z = 0; K3) = \chi(K3) = 24$$

and

$$\phi_{0,1}(\tau, z = 0) = 12$$

we obtain

$$\mathbf{EG}(\tau, z; T^4) = 0 \quad , \quad \mathcal{Z}(\tau, z; K3) = 2\phi_{0,1}(\tau, z) \ .$$

This clearly demonstrates the power of modularity in gaining extremely non-trivial information about the spectrum of $\mathcal{N} = (2, 2)$ SCFT.

Remark. The argument for the holomorphicity of the elliptic genus fails in an interesting way for theories whose spectrum has a continuous part. Due to the possible spectral asymmetry (i.e. non-perfect pairing between bosonic and fermionic states), the elliptic genus, when defined as a trace/integral over the full Hilbert space with continuous spectrum included, of such a theory could develop a non-trivial $\bar{q}$-dependence. For such an object the usual path-integral intuition still holds and the resulting non-holomorphic function transforms as a Jacobi form. Restricting to the discrete part of the spectrum, the analogous trace will be holomorphic but will no longer be modular. In particular, it will be a mock Jacobi form. As a result, in this context the holomorphic part of the elliptic genus is a well-defined notion both from a physical and mathematical point of view. From the physics perspective, the holomorphic part corresponds to the contribution from the discrete part of the spectrum [24, 27, 30, 59]. From the mathematical point of view, the holomorphic part corresponds to the holomorphic part of the harmonic Maass form [60]. We will encounter such a situation in §8.2.

Part II

Moonshine

In this part we describe classical and recent cases of moonshine connections, focussing on the mathematical statements and postponing the physics till the next part. We organise the different cases in terms of the weights of the modular objects involved, and we will see that in different weights the moonshine relations have features that are different in interesting ways.

4 Moonshine at weight zero

Here we review the two moonshine connections, monstrous and Conway moonshine, that occur at weight zero. They are the moonshine cases that are best understood at the moment, in terms of the specification of the modular objects, the origin of the symmetries, and their physical context.

4.1 Monstrous moonshine

Monstrous moonshine is arguably one of the most fascinating chapters of mathematics in the last century, where finite groups and modular objects were first noticed to be related via physical structures. As the theory of moonshine further develops, we believe that monstrous moonshine will remain the most distinguished example of moonshine phenomenon from various points of view. In this section we briefly describe the features of monstrous moonshine, and we refer to [40, 61, 62] and references therein for other excellent reviews of this beautiful story, in particular the historical aspects of it.

The term *monstrous moonshine*, coined in [63], refers to the unexpected connection between the representation theory of the Monster group $\mathbb{M}$ and the modular form

$$J(\tau) = \sum_{n\geq -1} a_n\, q^n = q^{-1} + 196884q + 21493760q^2 + 864299970q^3 + \cdots \ , \qquad (4.1)$$

which we encountered in (2.5). The development of monstrous moonshine was initiated with the key observation, due to McKay, that the coefficient 196884 in the q-expansion of J can be decomposed as $196884 = 1 + 196883$, where the summands are the dimensions of the two smallest irreducible represenations of $\mathbb{M}$. Similar decompositions were observed for the next few coefficients by Thompson in [64]:

$$\begin{aligned} 1 &= 1 \\ 196884 &= 196883 + 1 \\ 21493760 &= 21296876 + 196883 + 1 \\ 864299970 &= 842609326 + 21296876 + 2\cdot 196883 + 2\cdot 1 \\ &\cdots \end{aligned} \qquad (4.2)$$

where 1, 196883, 21296876, and 842609326 are dimensions of certain irreducible representations of $\mathbb{M}$. The observation led to the conjecture of the existence of an infinite-dimensional $\mathbb{Z}$-graded Monster module,

$$V = \bigoplus_{n\geq -1} V_n \ , \qquad (4.3)$$

such that $\dim V_n = a_n$ for all $n \geq -1$. In other words, J acquires the interpretation as the graded dimension of V

$$J(\tau) = \sum_{n\geq -1} \dim V_n q^n. \tag{4.4}$$

Notice that V_0 is empty, corresponding to the vanishing constant coefficient of J.

The conjecture. This conjecture as stated above is not interesting since one could take each V_n to contain $c(n)$ copies of the trivial representation of $\mathbb{M}$ to make (4.4) true, given the fact that all a_n are non-negative integers. To access the information on the $\mathbb{M}$-action, Thompson also proposed in [65] to look at the graded characters of V, the so-called *McKay-Thompson series* defined by

$$T_g(\tau) \coloneqq \sum_{n\geq -1} \mathrm{Tr}_{V_n}(g) q^n \ , \tag{4.5}$$

for each element $g \in \mathbb{M}$ (with $T_e = J$). Note that the q-series $T_g(\tau)$ must also have vanishing constant term. As is clear from the definition, the T_g are class functions, i.e. $T_g = T_{hgh^{-1}}$. As a result, there are at most 194 distinct T_g as $\mathbb{M}$ has 194 conjugacy classes. In fact, it turns out that T_g only gives rise to 171 distinct functions. The main point of monstrous moonshine lies in the fact that these graded trace functions also exhibit modular properties and are moreover the unique Hauptmoduls with no constant terms (cf. §2.1), as stated in the following astonishing conjecture made by Conway and Norton [63]:

Conjecture 4.1. (Monstrous Moonshine Conjecture)
For each $g \in \mathbb{M}$ the McKay-Thompson series T_g coincides the unique Hauptmodul J_{Γ_g} with expansion $q^{-1} + O(q)$ near $\tau \to i\infty$, for some genus zero subgroup $\Gamma_g \leq SL_2(\mathbb{R})$. Furthermore, each Γ_g contains $\Gamma_0(N)$ as a normal subgroup, for some N dividing the quantity $|g| \gcd(24, |g|)$.

Given the importance of this conjecture, we will pause to make a few comments. Note that Γ_g is often not a subgroup of $SL_2(\mathbb{Z})$; only for some g we have $\Gamma_g = \Gamma_0(N)$ (cf. (2.8)), for some N satisfying the conditions mentioned above. In general, Γ_g is a normaliser of $\Gamma_0(N)$ in $SL_2(\mathbb{R})$, which in general involves the so-called Atkin-Lehner involutions. For later purpose we will be particular interested in the groups of the form

$$\Gamma^{N+K} \coloneqq \left\{ \frac{1}{\sqrt{n}} \begin{pmatrix} an & b \\ cN & dn \end{pmatrix} \mid adn - bcN/n = 1,\ n \in K \right\}, \tag{4.6}$$

where $K < \mathrm{Ex}_N$ is a subgroup of the group of exact divisors of N. We say that e is an exact divisor of N if $e|N$ and $(f, \frac{N}{f}) = 1$, and they form a group with multiplication $f * f' = \frac{ff'}{(f,f')^2}$. An especially simple case is when N is a prime number p, and the full normaliser (corresponding to $K = \{1, p\}$) is given by

$$\Gamma_0(p)+ \coloneqq \left\langle \Gamma_0(p), \frac{1}{\sqrt{p}} \begin{pmatrix} 0 & -1 \\ p & 0 \end{pmatrix} \right\rangle . \tag{4.7}$$

A harbinger of monstrous moonshine, predating the observation by McKay, is the following observation made by Ogg [66]. He noted that $\Gamma_0(p)+$ defines a genus zero quotient on the upper-half plane if and only if

$$p \in \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 41, 47, 59, 71\}, \tag{4.8}$$

and this is precisely the set of primes dividing the order of the Monster group, and subsequently offered a bottle of Jack Daniel's to anyone who can explain the coincidence [66]. Monstrous moonshine sheds light on this mysterious coincidence through the fact that the Hauptmoduls of all the genus zero $\Gamma_0(p)+$ feature in moonshine as the McKay-Thompson series T_g for a $g \in \mathbb{M}$ with order p. In the case $\Gamma_g \not\subset SL_2(\mathbb{Z})$ the modularity of CFT does not help to explain the appearance of modularity for Γ_g, since in CFT modularity arises from the mapping class group of the torus (cf. §3). The crucial genus zero property of monstrous moonshine received a useful paraphrasing [67] as the property that these Hauptmoduls can be obtained as a *Rademacher sum*, a regularised sum over the images of the polar term (in this case q^{-1}) under the action of the appropriate subgroup of $SL_2(\mathbb{R})$ (in this case Γ_g). This *Rademacher summability* property subsequently played a key role in the discovery of umbral moonshine (cf. §5.2). Recently, the genus zero property is explained by noting that Γ_g plays the role of the stringy symmetry group in the string realisation of the Monster theory and by requiring a physical analyticity condition on the supersymmetric index of the theory. See §7 for more details.

The moonshine module. This conjecture was verified numerically by Atkin, Fong and Smith (cf. [68, 69]), following the idea of Thompson (see [40] for references). To be more specific, they showed that, for each $n \geq -1$, the q^n-coefficient of the Hauptmoduls specified in [63] coincide with the characters of a certain virtual representation of $\mathbb{M}$. A constructive verification was later obtained by Frenkel, Lepowsky and Meurman [70, 71], with the explicit construction of a Monster module $V = V^\natural$. This module has the structure of holomorphic VOA, i.e. a VOA with a single irreducible $V^\natural$-module, namely itself.

The starting point for constructing $V^\natural$ is 24 chiral bosons $X^i(z)$, compactified on the 24 dimensional torus $\mathbb{R}^{24}/\Lambda$ defined by the Leech lattice Λ. This results into a VOA $V(\Lambda)$ with central charge $c = 24$, leading to a partition function whose q-expansion starts with $Z_{V(\Lambda)}(\tau) = q^{-1} + \ldots$. This, together with the modular invariance, fixes the function to be the same as $J(\tau)$ up to an additive constant. At the same time, we know what this constant has to be since the Leech lattice has no root vectors and hence $\Theta_\Lambda(\tau) = \sum_{v\in\Lambda} q^{\langle v,v\rangle/2} = 1 + O(q^2)$, leading to

$$Z_{V(\Lambda)}(\tau) = \frac{\Theta_\Lambda(\tau)}{\eta^{24}(\tau)} = J(\tau) + 24 \ . \tag{4.9}$$

In other words, thanks to the root-free property of the Leech lattice the lattice vertex operators of the form $e^{ik\cdot\Phi}$ all have weight larger than one, and the only remaining weight one primaries are the 24 fields ∂X^i.

In order to have an exact matching with J we would like to remove these primaries, which can be achieved by a simple $\mathbb{Z}_2$ orbifold of $V(\Lambda)$, acting as $X^i \to -X^i$, which corresponds to the $\{\mathrm{id}, -\mathrm{id}\} \cong \mathbb{Z}_2$ symmetry of Λ, contained in $\mathrm{Aut}(\Lambda) \cong Co_0$. Indeed, one can easily compute the partition function of the orbifolded theory explicitly as follows. Note that the $\mathbb{Z}_2$-twined partition function of 24 chiral bosons is given by

$$Z\left(-\underset{+}{\square}\,;\tau\right) = \frac{1}{q\prod_{n>0}(1+q^n)^{24}} = \left(\frac{2\eta(\tau)}{\theta_2(\tau)}\right)^{12} . \tag{4.10}$$

The orbifold entails that $V^\natural$ is the direct sum of the $\mathbb{Z}_2$-invariant projections of the untwisted and twisted sectors respectively (cf. (3.30)). From (3.29) and (2.13) we have

$$\begin{aligned} Z_{V^\natural}(\tau) &= \frac{1}{2}\left(J(\tau) + 24 + \left(\frac{2\eta(\tau)}{\theta_2(\tau)}\right)^{12} - \left(\frac{2\eta(\tau)}{\theta_3(\tau)}\right)^{12} + \left(\frac{2\eta(\tau)}{\theta_4(\tau)}\right)^{12}\right) \\ &= J(\tau). \end{aligned} \tag{4.11}$$

It remains to see that $\mathrm{Aut}(V^\natural)$ is the Monster. Note that $\mathrm{Aut}V(\Lambda)$ has a continuous piece T which is a 24-dimensional torus corresponding to the translation symmetry of the chiral bosons and to the 24 weight-one primary fields ∂X^i. The total symmetry is captured by the (non-split) short exact sequence

$$1 \to T \to \mathrm{AutV}(\Lambda) \to Co_0 \to 1 \ . \tag{4.12}$$

The $\mathbb{Z}_2$-orbifold breaks the automorphism group to its discrete part $2^{24}.\,Co_0$, which preserves the decomposition $V^\natural = V^\natural_+ \oplus V^\natural_-$ and is suggestively similar to a certain maximal subgroup $2^{1+24}.\,Co_1$ of $\mathbb{M}$. It is clear from the contribution to the weight two (and similarly for weight three, four, ...) states in $V^\natural$ from $V^\natural_+$ and $V^\natural_-$ that the Monster must mix them and hence cannot preserve the (un)twisted sector individually. Note that the 196884-dimensionl space of weight two states of $V^\natural$ has the structure of a commutative and non-associative algebra (as is true for any VOA), and can be shown to be precisely the *Griess algebra* constructed in 1980 and used to construct the Monster group itself [72]. From this and the VOA structure of $V^\natural$ one can show that $\mathrm{Aut}(V^\natural)$ is indeed the Monster, and can be obtained by adjoining a certain order two symmetry mixing $V^\natural_\pm$ to the discrete part of $\mathrm{Aut}(V(\Lambda))$.

The proof of monstrous moonshine. To prove that the $V^\natural$ constructed by Frenkel, Lepowsky and Meurman indeed "does the job", one needs to show that

$$T_g^{V^\natural}(\tau) := \mathrm{Tr}_{V^\natural}\, g\, q^{L_0 - c/24} \tag{4.13}$$

coincides with the corresponding Hauptmodul J_{Γ_g} specified in [63]. It was known that the coefficients of Hauptmoduls satisfy certain recursive formulas and one can determine all coefficients from just a handful of them. In the simplest case the recursive formulas are encoded in the remarkable identity

$$p^{-1} \prod_{\substack{m>0 \\ n\in\mathbb{Z}}} (1 - p^m q^n)^{a_{mn}} = J(\rho) - J(\tau) \ , \tag{4.14}$$

independently discovered by Zagier, Borcherds and others. Here $p = e^{2\pi i \rho}$, and a_i denotes the q^i coefficients in the q-expansion of J (cf. (4.1)). This identity results in infinitely many relations between a_i, which enables one to completely fix all the coefficients from just a_1, a_2, a_3, a_5. Clearly, the proof can be achieved if one can show the existence of the same type of identities among the coefficients of $T_g^{V^\natural}(\tau)$, and just explicitly compare the handful of coefficients that are necessary to fix the whole functions on both sides.

This is precisely what Borcherds did, and he obtained the replication formulas by introducing the notion of a *generalised Kac-Moody algebra*, which can be viewed as a generalisation of Kac–Moody algebras that allows for imaginary simple roots. Subsequently, he constructed a generalised Kac-Moody algebra (also called "Borcherds-Kac-Moody algebra") $\mathfrak{m}$, called the Monster Lie algebra. Roughly speaking, the construction was achieved by studying the cohomology of a BRST-like operator, which acts on $V^\natural \times \Gamma^{1,1}$, where $\Gamma^{1,1}$ is the unique unimodular lattice of signature $(1,1)$. This construction has a natural interpretation in string theory of considering second quantised strings in the background of $V^\natural$. See §7 for more details.

Borcherds managed to derive the replication formulas (4.14) as the *denominator identities* of the Monster Lie algebra $\mathfrak{m}$ that he attached to $V^\natural$. As in usual Kac–Moody algebras, the denominator identity results from applying the Weyl-Kac character formula of a Lie algebra to the trivial representation, and in this case relates an infinite sum to an infinite product, precisely the structure we see in (4.14). Moreover, by considering the $\mathbb{M}$-action on $V^\natural$ one can also obtain from $\mathfrak{m}$ the analogous identity

$$p^{-1}\exp\left[-\sum_{k>0}\sum_{\substack{m>0\\ n\in\mathbb{Z}}} a^{g^k}_{mn}\frac{p^{mk}q^{nk}}{k}\right] = J_{\Gamma_g}(z) - J_{\Gamma_g}(\tau)\ , \tag{4.15}$$

satisfied by the other Hauptmoduls, where a^g_i are the q-expansion coefficients of J_{Γ_g}. Combining the above components then proves the validity of $V^\natural$ as the module of monstrous moonshine.

Generalised monstrous moonshine. In [73] Norton proposed a generalisation of monstrous moonshine under the name of *generalised monstrous moonshine*. He suggested that there is a rule to assign to each element $g \in \mathbb{M}$ a graded projective representation $V(g) = \bigoplus_{n\in\mathbb{Q}} V(g)_n$ of the centralizer group $C_{\mathbb{M}}(g)$, and to each pair (g,h) of commuting elements of $\mathbb{M}$ a holomorphic function $T_{(g,h)}$ on the upper half-plane $\mathbb{H}$, which satisfies the following conditions:

(i) $T_{(g^ah^c,g^bh^d)}(\tau) = \gamma\ T_{(g,h)}\left(\frac{a\tau+b}{c\tau+d}\right)$ with $\begin{pmatrix} a & b\\ c & d\end{pmatrix} \in SL_2(\mathbb{Z})$ and γ being a 24th root of unity.

(ii) $T_{(g,h)}(\tau) = T_{(k^{-1}gk,k^{-1}hk)}(\tau)$ with $k \in \mathbb{M}$.

(iii) There is a lift $\tilde{h}$ of h to a linear transformation on $V(g)$ such that

$$T_{(g,h)}(\tau) = \sum_{n\in\mathbb{Q}} \mathrm{Tr}_{V(g)_n}\left(\tilde{h}\ q^{n-1}\right)\ . \tag{4.16}$$

(iv) $T_{(g,h)}(\tau)$ is either a constant or a Hauptmodul for some genus-zero congruence subgroup of $SL_2(\mathbb{Z})$.

(v) $T_{(e,h)}(\tau)$ coincide with $T_h(\tau)$, the McKay-Thompson series attached to $h \in \mathbb{M}$ by monstrous moonshine.

As we can see from the discussion in §3.2, all of these properties, apart from (iv), can be understood in the framework of holomorpic orbifolds [74], applied to $V^\natural$. In particular, the function $T_{(g,h)}$ can be thought of the h-twined character of the twisted module $V^\natural_g$. The proof of generalised monstrous moonshine was carried out recently in [75], where a generalised Kac–Moody algeba $\mathfrak{m}_g$, generalising the monster Lie algebra $\mathfrak{m}$, is constructed for all $g \in \mathbb{M}$.

4.2 Conway moonshine

Conway moonshine establishes the relation between Co_0, related to Conway's sporadic group Co_1 by $Co_1 \cong \mathrm{Co}_0/\{\pm\mathrm{Id}\}$, and Hauptmoduls of certain genus zero subgroups of $SL_2(\mathbb{R})$.

Recall from §1.3 that Co_0 is isomaprhic to the automorphism group of the Leech lattice Λ. In this context, hints of Conway moonshine had already appeared in the original montrous moonshine paper [63], where the authors assigned genus zero groups $\Gamma_g < SL_2(\mathbb{R})$ to elements $g \in \mathrm{Co}_0$: let $\{\lambda_i, \lambda_i^{-1}\}_{i=1}^{12}$ be the 24 eigenvalues of the natural g-action on the Leech lattice $\Lambda \otimes_{\mathbb{Z}} \mathbb{C}$ (embedded in a complex vector space), then Γ_g is given by the invariance groups of the holomorphic function

$$t_g(\tau) := q^{-1} \prod_{n>0} \prod_{i=1}^{12} \left(1 - \lambda_i q^{2n-1}\right)\left(1 - \lambda_i^{-1} q^{2n-1}\right) = q^{-1} - \chi_g + \mathcal{O}(q) \ . \tag{4.17}$$

Note that $\chi_g = \sum_i (\lambda_i + \lambda_i^{-1})$ is generically non-vanishing, and t_g has non-zero constant terms unlike the monstrous moonshine functions discussed in the previous subsection.

Conway moonshine, on the other hand, introduces a Conway module $V^{s\natural}$ whose McKay–Thompson series coincide with Hauptmoduls with vanishing constant terms. It was developed in [76, 77] (see also [78], [79] and [80] for nice summaries of the construction). The Conway module $V^{s\natural}$ is the unique, up to isomorphisms, super-VOA (SVOA) with $c_{V^{s\natural}} = 12$ and $N = 1$ superconformal structure, characterized by the absence of states with weight 1/2. It can be constructed as a $\mathbb{Z}_2$ orbifold of the theory with eight bosons on the E_8 torus together with their fermionic superpartners. Alternatively, it can be constructed as a $\mathbb{Z}_2$ orbifold, acting as $k_a \to -k_a$, of 24 free chiral fermions k_a, $a = 1, 2, \dots, 24$. This is to be compared with the monstrous moonshine module $V^\natural$, where the corresponding Monster module $V^\natural$ is built as a $\mathbb{Z}_2$ orbifold of the Leech lattice VOA (24 chiral bosons compactified on $\mathbb{R}^{24}/\Lambda$), resulting in theory with $c_{V^\natural} = 24$ and no states of weight 1. It turns out that $V^{s\natural}$ has an interesting relation to stringy symmetries of K3 surfaces (see §8.3). In what follows we will give more details on the Conway module $V^{s\natural}$.

Consider 24 real chiral fermions k_a and the corresponding complex fermions

$$\psi_j^{\pm} = \frac{1}{\sqrt{2}} \left(k_{2j-1} \pm i k_{2j}\right) \ , \quad j = 1, \dots, 12 \ , \tag{4.18}$$

with the following non-vanishing OPEs and stress-energy tensor

$$\psi_i^{\pm}(z)\psi_j^{\mp}(w) \sim \frac{\delta_{ij}}{z-w} \ , \qquad L = -\frac{1}{2}\sum_{i=1}^{12} {:}\psi_i^+ \partial\psi_i^- + \psi_i^- \partial\psi_i^+{:} \ . \tag{4.19}$$

Denote by $\mathfrak{a}$ the 24-dimensional vector space spanned by the fermions. Since fermions allow for both periodic and anti-periodic boundary conditions, there exist two sectors in the theory. The antiperiodic (Neveu-Schwartz) sector contains a single ground state $|0\rangle$ and excitations of half-integer weight, while the periodic (Ramond) sector contains integral-weight excitations and has 2^{12} degenerate ground states. The degeneracy is due to the Clifford algebra satisfied by the zero modes,

$$\left\{\psi_{i,0}^{\pm}, \psi_{j,0}^{\pm}\right\} = 0 \ , \quad \left\{\psi_{i,0}^{\pm}, \psi_{j,0}^{\mp}\right\} = \delta_{ij} \ , \tag{4.20}$$

which moreover commute with L_0. As a result, one can build the Ramond ground states by acting with $\psi^-_{i,0}$ on a ground state $|s\rangle$ satisfying $\psi^+_{i,0}|s\rangle = 0$. Namely, the Ramond ground states are given by the mononomials

$$\psi^-_{i_1,0} \cdots \psi^-_{i_k,0}|s\rangle \ , \tag{4.21}$$

which form a spinor in twenty-four dimensions with Euclidean signature.

Next we want to construct an action of Co_0 on the states described above. To do so, recall that $\mathrm{Co}_0 \cong \mathrm{Aut}(\Lambda)$, so the Conway group is isomorphic to a subgroup of $SO(24)$ and we can make the natural identification $\mathfrak{a} = \Lambda \otimes_{\mathbb{Z}} \mathbb{C}$, i.e. let fermions "live" on the Leech lattice. Then consider a group element $g \in \mathrm{Co}_0$ with complex eigenvalues $\lambda_i^{\pm 1}$, and choose the basis of $\mathfrak{a}$ such that the fermions $\psi^\pm_i$ are acted upon as eigenvectors:

$$g\psi^\pm_i = \lambda_i^{\pm 1}\psi^\pm_i \ , \quad \lambda_i \equiv e^{2\pi i a_i} \ , \quad i = 1, \dots, 12 \ . \tag{4.22}$$

Moreover, since the ground states in the Ramond sector form a representation of the Clifford algebra associated to $\mathfrak{a}$, we should lift $G < SO(24)$ to a subgroup $\hat{G} < \mathrm{Spin}(\mathfrak{a})$. An element $x \in \mathrm{Spin}(\mathfrak{a})$ has the property $xux^{-1} \in \mathfrak{a}$ for $u \in \mathfrak{a}$. We define the *lift* $\hat{g} \in \hat{G} < \mathrm{Spin}(\mathfrak{a})$ of $g \in G < SO(\mathfrak{a})$ by requiring that it results in the same action as g when acting on $\mathfrak{a}$,

$$\hat{g}(u) := \hat{g}u\hat{g}^{-1} = gu \ , \quad \forall \ u \in \mathfrak{a} \ . \tag{4.23}$$

The map $u \mapsto \hat{g}(u)$ is a linear transformation on $\mathfrak{a}$ belonging to $SO(\mathfrak{a})$, so $\hat{g} \mapsto \hat{g}(\cdot)$ defines a map $\mathrm{Spin}(\mathfrak{a}) \to SO(\mathfrak{a})$ with kernel $\{\pm\mathbf{1}\}$, i.e. $\mathrm{Spin}(\mathfrak{a})$ is a double cover of $SO(\mathfrak{a})$. It turns out that for $G \cong \mathrm{Co}_0$ there exists a unique lift $\hat{G} \cong Co_0$ (see [77] for more details).

While the NS ground state $|0\rangle$ is invariant under Co_0, the group action on the 2^{12} Ramond ground states turns out to be

$$\hat{g}|s\rangle = \prod_{i=1}^{12} e^{\pi i a_i}|s\rangle = \nu|s\rangle \ , \quad \nu \equiv \prod_{i=1}^{12} \nu_i \ , \quad \nu_i \equiv e^{\pi i a_i} = \lambda_i^{1/2} \ , \tag{4.24}$$

where $|s\rangle$ is the ground state described in (4.21). Notice that a priori there is a sign ambiguity for ν_i, since it is the square root of λ_i. But actually the choice of sign is unique since the lift of Co_0 is unique. There is a further ambiguity in the definition of the g-action on the fermions, in that we can swap the complex eigenvalues. This translates into setting $\lambda_i \leftrightarrow -\lambda_i^{-1}$ in (4.22), and is referred to as a choice of polarisation.

The last step is to consider a $\mathbb{Z}_2 = \{1, \mathfrak{z}\}$ orbifold of the theory described so far, acting as $\mathfrak{z}\psi^\pm_i = -\psi^\pm_i$ on the fermions. In other words, it acts as $(-1)^F$ where F is the fermion number. Supposing that it acts trivially on both ground states $|0\rangle$ and $|s\rangle$ [77], it splits the two sectors into even/odd eigenspaces,

$$\mathrm{NS} = \mathrm{NS}^0 \oplus \mathrm{NS}^1 \ , \quad \mathrm{R} = \mathrm{R}^0 \oplus \mathrm{R}^1 \ , \tag{4.25}$$

where the eigenvalues of $\mathrm{NS}^j/\mathrm{R}^j$ are given by $(-1)^j$. From this point on, one can construct two closely related SVOAs. A useful description is by exploiting the fact that NS^0 forms a (bosonic) VOA on its own, that of the lattice D_{12}. Equivalently, it is the VOA associated to the affine Kac-Moody algebra $\widehat{\mathfrak{so}}(24)_1$, at level 1. The latter has four irreducible integrable modules, namely the vacuum module $A \cong \mathrm{NS}^0$, the vector module V, the spinor module S and the conjugate spinor module C. By extending the D_{12} VOA NS^0 by either of the spinor modules, one arrives at two SVOAs,

$$\begin{aligned} V^{f\natural} &= \mathrm{NS}^0 \oplus S \cong \mathrm{NS}^0 \oplus \mathrm{R}^0 \\ V^{s\natural} &= \mathrm{NS}^0 \oplus C \cong \mathrm{NS}^0 \oplus \mathrm{R}^1 \ . \end{aligned} \tag{4.26}$$

From the orbifold point of view we have the identifications $S \cong \mathrm{R}^0$ and $C \cong \mathrm{R}^1$ in our notation.

The two SVOAs $V^{f\natural}$ and $V^{s\natural}$ are isomorphic as SVOAs, and are uniquely characterized by their central charge $c_{V^{f\natural}} = c_{V^{s\natural}} = 12$ and the absence of weight 1/2 states. In [76] is was shown that the $N = 1$ supercurrent of $V^{f\natural}$ is fixed by a subgroup of Spin(24) isomorphic to Co_0, which is identified by the group $\hat{G}$ in the notation above. In particular, note that $Z(\mathrm{Spin}(24)) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ where the first $\mathbb{Z}_2$ factor can be identified with the kernel of Spin(24) → $SO(24)$ and the latter with the center of $SO(24)$. The centre $Z(\hat{G}) \cong \mathbb{Z}_2$ can be identified with the second $\mathbb{Z}_2$ in Z(Spin(24)), and has the same action as the $\mathbb{Z}_2$ in the orbifold construction. As a result, it follows immediately from (4.26) that Co_0 does not act faithfully on $V^{f\natural}$, since the latter is invariant under the action of the centre $Z(\hat{G})$. Instead, $V^{f\natural}$ carries a faithful action of the quotient group $\hat{G}/\mathbb{Z}_2 \cong Co_1$. On the other hand, Co_0 acts faithfully on $V^{s\natural}$, and this is ultimately the reason why we consider $V^{s\natural}$ instead of $V^{f\natural}$ in what follows. Another notable difference between $V^{f\natural}$ and $V^{s\natural}$ is that the $N = 1$ supercurrent in $V^{f\natural}$ fixed by Co_0 is not contained in $V^{s\natural}$, but rather in $V^{s\natural}_{\mathrm{tw}}$ (inside the R^0 part). A "canonically twisted" (or Ramond sector) module for $V^{s\natural}$ can also be constructed as

$$V^{s\natural}_{\mathrm{tw}} = \mathrm{NS}^1 \oplus \mathrm{R}^0 \ , \tag{4.27}$$

which is twisted with respect to the $Z(\hat{G})$ symmetry. The action of Co_0 on this twisted module is also faithful.

In order to formulate the Conway moonshine statement, first define the functions

$$\begin{aligned} \eta_{\pm g}(\tau) &:= q \prod_{n=1}^{\infty} \prod_{i=1}^{12} \left(1 \mp \lambda_i^{-1} q^n\right)\left(1 \mp \lambda_i q^n\right) \\ C_{\pm g} &:= \nu \prod_{i=1}^{12} \left(1 \mp \lambda_i^{-1}\right) = \prod_{i=1}^{12} \left(\nu_i \mp \nu_i^{-1}\right) \ . \end{aligned} \tag{4.28}$$

The twined partition functions of Conway moonshine are then given by

$$\begin{aligned} T^s_g(\tau) &:= \mathrm{str}_{V^{s\natural}} \left[\hat{g} q^{L_0 - 1/2}\right] = \mathrm{tr}_{V^{s\natural}} \left[\mathfrak{z} \hat{g} q^{L_0 - 1/2}\right] = \frac{\eta_g(\tau/2)}{\eta_g(\tau)} + \chi_g \\ T^s_{g,\mathrm{tw}}(\tau) &:= \mathrm{str}_{V^{s\natural}_{\mathrm{tw}}} \left[\hat{g} q^{L_0 - 1/2}\right] = \mathrm{tr}_{V^{s\natural}_{\mathrm{tw}}} \left[\mathfrak{z} \hat{g} q^{L_0 - 1/2}\right] = C_g \eta_g(\tau) - \chi_g \ , \end{aligned} \tag{4.29}$$

where the super-gradings can be defined by inserting $(-1)^F$ into the trace, whose action coincides with that of $\mathfrak{z}$ (recall that we identified $\mathbb{Z}_2 = \{1, \mathfrak{z}\}$ with the centre of Co_0). The main theorem of Conway moonshine states [77]:

Theorem 4.2. *The function $T^s_g(2\tau) = q^{-1} + O(q)$ is a Hauptmodul for a genus zero group $\Gamma_g < SL_2(\mathbb{R})$ that contains some $\Gamma_0(N)$, for every $g \in Co_0$. If g has a fixed point in its action on Λ, then $T^s_{g,tw}(\tau)$ is equal to the constant $-\chi_g$. Furthermore, if g has no such fixed point, then $T^s_{g,tw}(\tau)$ is also a Hauptmodul for a genus zero* *subgroup of* $SL_2(\mathbb{R})$.

In §8.3 we will see a prominent role played by Conway moonshine in relation to the symmetries of $K3$ CFTs.

5 Moonshine at weight one-half

Somewhat unexpectedly, a wave of moonshine development started in 2010 which led to the discovery of many more examples of moonshine connections. The modern examples share some similarities, but also display important differences with the classical moonshine examples discussed in §4. The modular objects in these examples are typically mock modular forms, an important and natural generalisation of modular forms introduced in §2.3, which furthermore have non-vanishing weights. The first and very fruitful arena that was explored is that of weight 1/2 mock modular forms. In this section we will describe interesting examples of moonshine relating finite groups and weight 1/2 mock modular forms.

5.1 Mathieu moonshine

The first example of the new type of moonshine, *Mathieu moonshine*, was initiated with certain observations about the weight 1/2 mock modular form H introduced in (2.42), in an analogous fashion as how observations about the classical J function initiated the development of monstrous moonshine. In [81] it was pointed out that the first few Fourier coefficients of H coincide with twice the dimensions of certain irreducible representations of the largest sporadic group M_{24}. Moreover, this observation was made in an interesting physical context which we will mention briefly below, and will explain in more detail in §8.

By now we have understood that, from many different points of view, Mathieu moonshine should really be thought of as a component of umbral moonshine, which we will review in the next subsection. However, in many ways Mathieu moonshine stands out among the other cases of umbral moonshine, not just historically but also in terms of its direct relation to the $K3$ elliptic genus. As a result, we will devote a separate subsection to Mathieu moonshine before discussing umbral moonshine.

Recall that the mock modular form H (2.42) can be viewed as arising from a meromorphic Jacobi form ψ given in (2.45). Using the relation between ψ and the $K3$ elliptic genus (2.44), as well as the identity $(\theta_{2,1}-\theta_{2,-1})(\tau,z) = -i\theta_1(\tau,2z)$, we obtain the following relation between the elliptic genus of $K3$ (cf. §3.3) and the mock modular form H:

$$\mathbf{EG}(\tau,z;K3) = \frac{\theta_1^2(\tau,z)}{\eta^3(\tau)}\left(24\,\mu(\tau,z)+H(\tau)\right)\ , \tag{5.1}$$

where

$$\mu(\tau,z) = -i\frac{\theta_1^2(\tau,z)}{\theta_1(\tau,2z)\eta^3(\tau)}\mathrm{Av}^{(2)}\left[\frac{y+1}{y-1}\right] = \frac{-iy^{1/2}}{\theta_1(\tau,z)}\sum_{\ell=-\infty}^{\infty}\frac{(-1)^\ell y^\ell q^{\ell(\ell+1)/2}}{1-yq^\ell}\ . \tag{5.2}$$

Note that while none of the two summands at the right-hand side of (5.1) transforms modularly, their modular anomalies cancel and the left-hand side is a perfectly well-behaved Jacobi form, as discussed in §2 and §3.3. In particular, a simple way to derive the shadow of H is by studying the modular properties of the Appell–Lerch sum $\mu(\tau,z)$ [21,82]. We will see in §8.1-8.2 the two interesting physical interpretations of the above splitting (5.1) of $\mathbf{EG}(K3)$, one in terms of the characters of $\mathcal{N}=4$ superconfomal algebra and one in terms of the elliptic genus of du Val singularities.

The aforementioned observation on the conspicuous relation between the first few coefficients of the mock modular form H and certain representations of M_{24} led to the suspicion that there exists a $\mathbb{Z}$-graded, infinite-dimensional M_{24}-module $K=\bigoplus_{n=1}^{\infty}K_n$ underlying H, namely $H(\tau)=q^{-\frac{1}{8}}\big(-2+\sum_{n=1}^{\infty}q^n\,(\dim(K_n)\big)$. A natural question is thus whether the other summand in the splitting of the Jacobi form $\mathbf{EG}(K3)$ (5.1) harbors an action of M_{24} as well. A simple guess arises from the fact that M_{24} is a subgroup of the permutation group S_{24} and as a result has a defining permutation representation $\mathbf{R}$, of dimension 24. A natural proposal for the "twined" version of (5.1) is therefore

$$\phi_g(\tau,z)=\frac{\theta_1^2(\tau,z)}{\eta^3(\tau)}\left((\mathrm{Tr}_{\mathbf{R}}\,g)\,\mu(\tau,z)+H_g(\tau)\right)\ , \tag{5.3}$$

where H_g denotes the graded characters of the M_{24} module K. Following the spirit of monstrous moonshine, we say that there is a non-trivial moonshine connection if all such ϕ_g transform nicely as Jacobi forms under some $\Gamma_g\subseteq SL_2(\mathbb{Z})$. Physical considerations reviewed in §3.2 moreover suggest that $\Gamma_0(|g|)\subseteq\Gamma_g$.

Fortunately, the possibility for this type of Jacobi forms is very limited and we are constrained to consider

$$\phi=c\,\phi_{0,1}+F\,\phi_{-2,1}\ , \tag{5.4}$$

where $\phi_{0,1}$ and $\phi_{-2,1}$ are given in (2.31), $c\in\mathbb{C}$, and F is a weight two modular form for Γ_g, possibly with a non-trivial multiplier system when $c=0$. The dimension of the space of possible F is often small for the $SL_2(\mathbb{Z})$-subgroup Γ_g we are interested in. For instance, when $\Gamma_g=SL_2(\mathbb{Z})$ the only possible weight two form is $F=0$. Hence, knowing the first few of the Fourier coefficients of ϕ_g, dictated by our guesses for the first few M_{24}-representations, is often sufficient to fix the whole function. As a result, not long after the original observation [81], candidates for the McKay–Thompson series were proposed for all conjugacy classes $[g]\subset M_{24}$ in [83–86], and they take the form

$$\phi'_g(\tau,z)=\frac{\mathrm{Tr}_{\mathbf{R}}\,g}{12}\,\phi_{0,1}(\tau,z)+\tilde{T}_g(\tau)\,\phi_{-2,1}(\tau,z)\ , \tag{5.5}$$

where the functions $\tilde{T}_g(\tau)$ are weight 2 modular forms explicitly specified in the references given above and collected in Table 2 of [87]. More precisely, these ϕ_g for any $g\in M_{24}$ are weak Jacobi form of weight zero and index one satisfying the elliptic invariance $\phi_g|_1(\lambda,\mu)=\phi_g$ for all $(\lambda,\mu)\in\mathbb{Z}^2$ (cf. (2.22)), and transform as

$$\phi'_g(\tau,z)=\rho_{n_g|h_g}(\gamma)\,e\left(-\frac{cz^2}{c\tau+d}\right)\phi'_g\big(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\big)\ ,$$

for $\gamma\in\Gamma_0(|g|)$, where the multiplier $\rho_{n_g|h_g}$ is summarised in [87].

In terms of the weak Jacobi forms (5.5), the main statement of Mathieu moonshine is the following.

Conjecture 5.1. *There exists a naturally defined $\mathbb{Z}$-graded, infinite-dimensional M_{24} module $K=\bigoplus_{n=1}^{\infty}K_n$ such that for any $g\in M_{24}$, the graded character*

$$H_g(\tau):=q^{-\frac{1}{8}}\big(-2+\sum_{n=1}^{\infty}q^n\,(\mathrm{Tr}_{K_n}\,g)\big) \tag{5.6}$$

satisfies $\phi_g = \phi'_g$, *where* ϕ_g *is as defined in* (5.3) *and* ϕ'_g *is the explicitly given weak Jacobi form* (5.5). *Moreover, the representations* K_n *are even in the sense that they can all be written in the form* $K_n = k_n \oplus k_n^*$ *for some* M_{24}*-representations* k_n *and their dual representation* k_n^*.

A proof of the key fact in the above conjecture, namely the existence of an M_{24}-module $K = \bigoplus_{n=1}^{\infty} K_n$ such that (5.8) holds, has been attained in [88]. However, a construction of the module K, analogous to the construction of $V^\natural$ by Frenkel–Lepowsky–Meurman in the case of monstrous moonshine, is still absent. Therefore in no way do we know why K should be "natural". As explicit data, the first few Fourier coefficients of the q-series $H_g(\tau)$ and the corresponding M_{24}-representations are given in [87].

Note that the above implies that there is a M_{24}-supermodule underlying all terms in the q-, y-expansion of the $K3$ elliptic genus. It is hence tempting to endow the McKay–Thompsen series ϕ_g with the physical interpretation as twined elliptic genera of $K3$ CFT (cf. (3.35)). This (im)possibility will be extensively discussed in §8. Finally, note that the modular form property of ϕ_g, as well as the mock modular property of $\mu(\tau, z)$, immediately lead to the fact that H_g are also mock modular forms. Explicitly, they are given by

$$H_g(\tau) = \frac{\mathrm{Tr}_{\mathbf{R}}\, g}{24} H(\tau) - \frac{\tilde{T}_g(\tau)}{\eta(\tau)^3} \,, \tag{5.7}$$

and they are weight 1/2 mock modular forms with shadows given by $(\mathrm{Tr}_{\mathbf{R}}\, g)\eta^3(\tau)$, generalising the mock modular property of $H(\tau)$ discussed around (2.43).

5.2 Umbral moonshine

A few years after the discovery of Mathieu moonshine, it was realised that it is in fact just one instance of a larger system of moonshine, called "umbral moonshine" [89,90]. There are in total 23 instances of umbral moonshine, which admit a uniform description (see Figure 2). The main statement of umbral moonshine is as follows.

Conjecture 5.2. *Let* G^X *be one of the 23 finite groups specified in* (5.9), m *be the corresponding positive integer specified in* (5.14), *and* I^X *be the specific subset of* $\{1, 2, \ldots, m-1\}$ *described in* (5.15). *Then there exists a naturally defined bi-graded, infinite-dimensional* G^X*-module*

$$K^X = \bigoplus_{r \in I^X} \bigoplus_{\substack{D \le 0 \\ D = r^2 \bmod 4m}} K^X_{r,D}$$

such that for any $g \in G^X$ *and for any* $r \in I^X$, *the graded character ("corrected" by a polar term* $-2q^{-\frac{1}{4m}}$ *as in below) coincides with the component* $H^X_{g,r}$ *of a vector-valued mock modular forms* $H^X_g = (H^X_{g,r})_{r \in I^X}$:

$$H^X_{g,r} = -2q^{-\frac{1}{4m}}\delta_{r,1} \; + \sum_{\substack{D \le 0 \\ D = r^2 \bmod 4m}}^{\infty} q^{-D/4m}\, (\mathrm{Tr}_{K^X_{r,D}}\, g). \tag{5.8}$$

In what follows we will briefly describe the specification of the main players, the finite groups G^X and the mock modular forms H^X_g, in the above conjecture. See Figure 2.

The starting point of this uniform construction are the 23 Niemeier lattices N^X introduced in §1.3. Recall that they are uniquely labelled by their root systems. We will denote by X the root systems, by N^X the corresponding Niemeier lattices, and by G^X the finite groups arising from the automorphisms $\mathrm{Aut}(N^X)$ via (1.13):

$$G^X := \mathrm{Aut}(N^X)/\mathrm{Weyl}(X) \tag{5.9}$$

These are the finite groups relevant for umbral moonshine and we will refer to them as the *umbral groups*.

On the modular side, we use the root system X to specify certain mock modular forms related to the finite group G^X. To explain how this is done, first recall that the McKay–Thompson series T_g in monstrous moonshine and the mock modular forms H_g in Mathieu moonshine have very special properties. First, once their (mock) modular data (consisting of the group Γ_g, the weight, and the multiplier) are specified, the functions are completely determined by the analyticity property of how they grow near the cusps $i\infty \cup \mathbb{Q}$. Second, they have "optimal growth" in the following sense. These functions 1) are bounded at all cusps that are not Γ_g-equivalent to $i\infty$ and 2) have the slowest possible growth near $i\infty$ that is compatible with the modular data. For instance, in the case of monstrous moonshine it is elementary to see that a modular form satisfying condition 1) for $\Gamma_g \supset \langle T\rangle$ must behave like $q^{-n}(1+O(q))$ for some integer n near the cusp $i\infty$. As a result, the condition 2) states that $n = 1$, which is indeed the case for the moonshine functions T_g. Another way to state the above is to say that the functions in monstrous and Mathieu moonshine can be written in terms of Rademacher sums over the minimal polar term in the expansion near $i\infty$ [67, 91]. See also §4.1 for a discussion on Rademacher sums.

The functions of umbral moonshine turn out to have analogous uniqueness properties, and the relevant concept here is the notion of optimal mock Jacobi forms. We will first focus on the case $g = e$ and $\Gamma_g = SL_2(\mathbb{Z})$. Let $\psi = \sum_r h_r\theta_{m,r}$ be a mock Jacobi form of weight one and index m. We say it is an optimal mock Jacobi form if

$$h_r(\tau) = O(q^{-\frac{1}{4m}}) \tag{5.10}$$

as $\Im(\tau) \to \infty$, for each $r \in \mathbb{Z}/2m$. For instance, the function ψ^{3E_8} defined in §2.3 is an optimal mock Jacobi since it has index 30 and $H_1^{3E_8}(\tau) = -2q^{-\frac{1}{120}}(1+O(q))$, while $H_7^{3E_8}$ vanishes at $\Im(\tau) \to \infty$ (cf. (2.38)). Similarly, ψ^{24A_1} is an index 2 optimal mock Jacobi form.

At weight one, the space of such optimal mock Jacobi forms turns out to be very restricted: the mock modular transformation property together with the pole structure of the functions near the cusps are sufficient to determine the whole q-series. In particular, they can be obtained as simple Rademacher sums involving only the polar parts as input. Such forms are even more scarce if we want them to have non-transcendental Fourier coefficients. Note that this must be the case for the function to play a role in moonshine, since the graded dimensions are necessarily integers and of course non-transcendental. In [15] it is shown that if ψ is such a form, it must lie in a 34-dimensional space, irrespective of its index. Moreover, inside this 34-dimensional space there are 39 special elements (which span the 34-dimensional space) distinguished by their special symmetries. Recall that Atkin–Lehner symmetries normalising $\Gamma_0(m)$ are specified by a subgroup K of the group of exact divisors Ex_m (cf. (4.6)). Given such a pair m and K, we say that an index m mock Jacobi form ψ is *K-symmetric* if

$$\psi = \sum_{r \bmod 2m} h_r \theta_{m,r} = \sum_{r \bmod 2m} h_r \theta_{m,a(n)r} \quad \text{for all } n \in K \ , \tag{5.11}$$

where, for a given n, we define $a(n)$ to be the unique element in $\mathbb{Z}/2m$ satisfying

$$a(n) = \begin{cases} 1 \bmod 2m/n \\ -1 \bmod 2n \end{cases} .$$

Note that the symmetry is an involution, since $a^2 = 1 \bmod 2m$. For instance, the mock Jacobi form ψ^{3E_8} introduced in (2.38) is invariant under the action of $K = \{1, 6, 10, 15\} < \mathrm{Ex}_{30}$, corresponding to $a(n) = 1, 11, 19, 29$. The surprising result in [15] then states that a non-vanishing K-symmetric index m optimal mock Jacobi form at weight one has non-transcendental coefficients if and only if the corresponding $SL_2(\mathbb{R})$ subgroup Γ^{m+K} defines a genus zero quotient in the upper-half plane. Recall that these genus zero groups also play an important role in monstrous moonshine. Note that we necessarily need to have $m \notin K$ (referred to as the "non-Fricke" property) for the mock Jacobi form to be non-vanishing, since at weight one has $\sum_r h_r \theta_{m,r} = -\sum_r h_r \theta_{m,-r}$ and $a(m) = -1$. There are just 39 such non-Fricke genus zero groups $\Gamma^{m+K} < SL_2(\mathbb{R})$ and we will denote the corresponding unique optimal mock Jacobi form, with the normalisation

$$h_1 = -2q^{-\frac{1}{4m}}(1 + O(q)) \ , \tag{5.12}$$

by ψ^{m+K}. In fact, these 39 distinguished optimal mock Jacobi forms ψ^{m+K} turn out to have Fourier coefficients that are not only non-transcendental, but also integral. Moreover, 23 among the 39 have positive coefficients in the following sense. By writing (cf. (2.28))

$$\psi = \sum_{1 \leq r \leq m-1} h_r \hat{\theta}_{m,r} \ ,$$

h_r has the expansion

$$h_r = \begin{cases} -2q^{-1/4m} + \sum_{n \geq 0} c_{r,n} q^{n/4m}, & \text{if } r^2 = 1 \bmod 4m \\ \sum_{n \geq 0} c_{r,n} q^{n/4m}, & \text{otherwise} \end{cases} , \tag{5.13}$$

with $c_{r,n} \in \mathbb{Z}_{\geq 0}$. This positivity property makes it possible for it to be the graded dimensions of finite group representations[3]. To sum up, for any index, there are 23 special mock Jacobi forms of weight one for $SL_2(\mathbb{Z})$ distinguished by

1. the optimality growth condition (5.10),
2. the Atkin–Lehner symmetries (5.11),
3. the normalisation (5.12),
4. the positivity and integrality of the coefficients (5.13).

---

[3] It is believed [15] that the remaining 16 optimal mock Jacobi forms ψ^{m+K} with positive and negative integral coefficients have also umbral type moonshine attached to them, but with additional supermodule structure that accounts for the minus sign.

The interesting observation is that these ψ^{m+K} with positivity properties are in one-to-one correspondence with the 23 Niemeier root systems X! To explain this correspondence, first recall the ADE classification of the modular invariant combination of $\widehat{A_1}$ affine Lie algebras [47], which has led to a classification of $\mathcal{N}=2$ superconformal minimal models with spectral flow symmetries. Their classification gives rise to a square matrix Ω^Y of size $2m$ for each simply-laced root system Y, where m coincides with the Coxeter number of Y. Moreover, the term $(\Omega^Y_{r,r}-\Omega^Y_{r,-r})$ coincides with the multiplicity of r as a Coxeter exponents (the degrees of the invariant polynomials shifted by one) of Y and takes values in $\{0,1,2\}$. The above can be generalised to a union of simply-laced root systems with the same Coxeter number (recall that this is indeed the case for Niemeier root systems) $X=\cup_i Y_i$ by defining $\Omega^X=\sum_i \Omega^{Y_i}$. Then the mock Jacobi form $\psi^X=\psi^{m+K}$, with theta-decomposition $\psi^X=\sum_r H^X_r\theta_{m,r}$, corresponding to the Niemeier root system X, display the following relations to X.

1. The Coxeter number of X coincides with the index of the ψ^X,

$$m=\mathrm{Cox}(X)\ . \tag{5.14}$$

2. The matrix Ω^X and the form $\psi^X=\psi^{m+K}$ have the same Atkin–Lehner symmetries: $(\Omega^X)_{r,r'}=(\Omega^X)_{r,a(n)r'}$ for all $n\in K$. Using these symmetries, it is convenient to define a set I^X of the orbits of the Atkin–Lehner symmetry group acting on $\{1,\dots,m-1\}$ (in $\mathbb{Z}/2m$), labelling the independent components H^X_r of the vector-valued mock modular form (H^X_r) and leading to

$$\psi^X=\sum_{r\in I^X}H^X_r\sum_{n\in K}\hat{\theta}_{m,a(n)r}\ . \tag{5.15}$$

3. The shadow of ψ^X is determined by Ω^X. More precisely, the completion of ψ^X is specified by the skew-holomorphic Jacobi form $\sigma=\sum_r\overline{\theta^1_{m,r}}\Omega^X_{r,r'}\theta_{m,r'}$ (cf. (2.37) and the preceeding text).

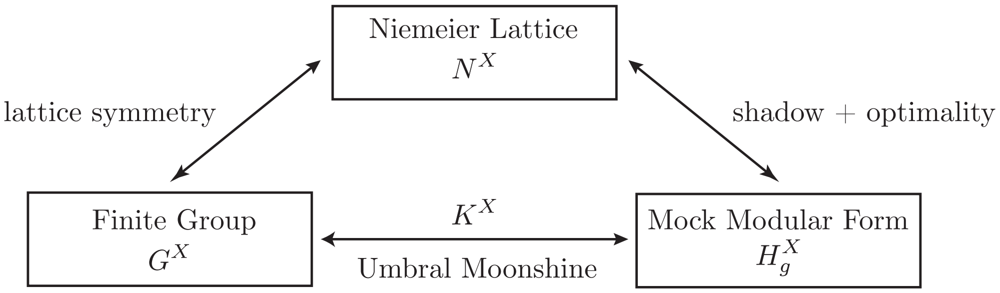
The description of ψ^{24A_1} and ψ^{3E_8} in §2.3 provides examples of the above. The mock Jacobi form ψ^X then gives us the vector-valued mock modular forms $H^X=(H^X_r)$ which will play the role of graded dimensions of the module for the umbral group G^X. In other words, we have $H^X_r=H^X_{e,r}$ in Conjecture 5.2. For the case $X=24A_1$ this is the Mathieu moonshine function ψ^{24A_1} we discussed in §2 and §5.1. Another simple example is when for $X=12A_2$, where $\psi^X=\sum_{r=1,2}H^X_r\hat{\theta}_{3,r}$ (so $m=3$ and $I^X=\{1,2\}$), with

$$\begin{aligned}H^X_1(\tau)&=2q^{-1/12}(-1+16\,q+55\,q^2+144\,q^3+\dots)\\H^X_2(\tau)&=2q^{8/12}(10+44\,q+110\,q^2+\dots).\end{aligned} \tag{5.16}$$

At the same time, the symmetries of the corresponding Niemeier lattice gives $G^X\cong 2.M_{12}$. The relation between the finite group G^X and the vector-valued mock modular form H^X can be observed from the fact that the group $2.M_{12}$ has irreducible representations of dimensions $16,55,144$ as well as $10,44,110$.

After specifying the mock Jacobi forms for $SL_2(\mathbb{Z})$, in order to describe the moonshine relation we also need a set of mock Jacobi forms $\psi^X_g=\sum_r H^X_{g,r}\theta_{m,r}$, one for each conjugacy class $[g]\subset G^X$, for subgroups of $SL_2(\mathbb{Z})$. The mock modular forms $H^X_g=(H^X_{g,r})$ will then play the role of graded characters of the umbral moonshine module, as described in Conjecture 5.2. This can be achieved in a way

Figure 2: The construction of umbral moonshine.



largely analogous to the $SL_2(\mathbb{Z})$ case, though additional subtleties do emerge and extra care needs to be taken. We refer to [92] for more details.

Once the mock modular forms H_g^X are specified, it is trivial to verify the existence of the G^X-module $K_{r,D}^X$ in Conjecture 5.2 term by term, namely one D at a time. Moreover, the existence of the whole umbral module $K^X = \bigoplus_r \bigoplus_D K_{r,D}^X$ has been proven mathematically using properties of (mock) modular forms [88,93]. However, the construction, or even an understanding of the exact nature of K^X, is not yet obtained in general. Construction of K^X has so far only been achieved for certain particularly simple cases of umbral moonshine, corresponding to Niemeier root systems $3E_8$ [94], $4A_6$ and $2A_{12}$ [95], $4D_6$, $3D_8$, $2D_{12}$ and D_{24} [96], as well as $6D_4$ [97]. The construction in [94] relies on special identities satisfied by the mock modular forms $H_{g,r}^{3E_8}$ relating it to a lattice-type sum, while in [95,96] the modules are constructed using the interpretation of the meromorphic Jacobi forms associated to Ψ_g^X as the twined partition function of certain vertex operator algebras (or chiral CFTs). In [97] the module for the $6D_4$ case of umbral moonshine is constructed by exploiting the relation between the (twined) $K3$ elliptic genus, umbral and Conway moonshine, which we will explain in §8.2. Moreover, generalised umbral moonshine, analogous to the generalised monstrous moonshine discussed in §4.1, has been established in [98], hinting that some elements of CFT/modular tensor category structure should be present at the umbral moonshine module K^X. Despite these results, it is fair to say that a uniform construction of the umbral module, reflecting the uniform description of umbral moonshine, is currently one of the biggest challenges in the study of moonshine.

We will end our review on umbral moonshine by noting a special property, called *discriminant property*, of umbral moonshine. It relates the discriminants D (the power of individual terms $q^{-D/4m}$ in the q-series H_r^X) and the number field generated by the characters of representations showing up in $K_{r,D}^X$. For instance, in the case $X = 24A_1$, the M_{24}-representation underlying the $q^{7/8}$ term in $H_g = H_{g,1}^{24A_1}$ is $K_{-7}^{24A_1} = \rho \oplus \rho^*$, where ρ is a 45-dimensional irreducible representation and ρ^* is the dual representation. At the same time, $\mathrm{Tr}_\rho\, g$ (and hence also $\mathrm{Tr}_{\rho^*}\, g$) generates the field $\mathbb{Q}(\sqrt{-7})$. Analogous relations continue for larger q-power as long as $\mathbb{Q}(D) = \mathbb{Q}(-7)$, and similar properties hold uniformly for all 23 cases of umbral moonshine. At present there is no physical understanding of this surprising and profound-looking property.

5.3 Thompson moonshine

The Thompson sporadic group Th is a subgroup of the Monster that has order $|Th| \sim 9 \cdot 10^{16}$ and contains 48 conjugacy classes. Thompson moonshine describes a special relation between the representation theory of Th and certain weakly holomorphic modular forms of weight 1/2. It was first discussed in [99] and is the first instance of the so-called *skew-holomorphic moonshine* [100] relating skew-holomorphic Jacobi forms and finite group representations. See §2.2 and §9 for more mathematical and physical information on skew-holomorphic Jacobi forms.

Just like Mathieu and umbral moonshine, Thompson moonshine relates finite groups connected to the Monster (cf. §1.3) and weight 1/2 modular objects. However, there are a few differences. First, the modular objects in Thompson moonshine are modular while in umbral moonshine they are mock modular with non-vanishing shadows. Second, there is a super-structure in the Thompson moonshine module, i.e. the representations come with signs. See Conjecture 5.3. That said, this extra $\mathbb{Z}_2$-grading structure is very simple and completely controlled by m mod 4, where m is the natural $\mathbb{Z}$-grading corresponding to the exponents of q in the moonshine function. Finally, we remark that Thompson moonshine also enjoys a discriminant property analogous to that of umbral moonshine.

To describe Thompson moonshine, we start with the following conjecture [99]:

Conjecture 5.3. *There exists a $\mathbb{Z}$-graded Th-supermodule (cf. (1.20))*

$$W = \bigoplus_{\substack{m \geq -3 \\ m \equiv 0,1 \bmod 4}}^{\infty} W_m \tag{5.17}$$

such that the graded dimensions of W_m are related to a certain weight 1/2 weakly holomorphic modular form $\mathcal{F}_3$ (5.20) by

$$\mathcal{F}_3(\tau) = \sum_{\substack{m=-3 \\ m \equiv 0,1 \bmod 4}} \mathrm{Str}_{W_m}(1) q^m \ . \tag{5.18}$$

Furthermore, the associated McKay-Thompson series

$$\mathcal{F}_{3,g}(\tau) = \sum_{\substack{m=-3 \\ m \equiv 0,1 \bmod 4}} \mathrm{Str}_{W_m}(g) q^m \ , \tag{5.19}$$

are given by certain weakly holomorphic weight 1/2 modular forms, with $\mathcal{F}_{3,e} = \mathcal{F}_3$. Finally, W_m is even under the superspace $\mathbb{Z}_2$-grading if $m \geq 0$ and $m \equiv 0 \bmod 4$ or when $m = -3$, and is odd otherwise.

Let's discuss the elements of this conjecture. The function $\mathcal{F}_3$ is defined by

$$\mathcal{F}_3(\tau) := 2f_3(\tau) + 248\theta(\tau) \ , \tag{5.20}$$

where $\theta(\tau)$ is the Jacobi theta function discussed in §2.1. We next describe the specification of f_3.

Define the Kohnen-plus space $M^+_{1/2}$ as the set of functions that transform like θ, and additionally satisfy the condition $c(n) = 0 \ \forall \ n \neq 0, 1 \bmod 4$ in their Fourier expansions $\sum c(n)q^n$. It can be shown that in fact $M^+_{1/2}$ is one-dimensional, spanned by θ itself. However, if we allow for poles at the cusps, we can get a nontrivial space

$M^{!,+}_{1/2}$ of weakly holomorphic modular forms of weight 1/2 in the Kohnen-plus space. In [101], Zagier constructed an infinite-dimensional basis for $M^{!,+}_{1/2}$, spanned by the (unique) functions

$$f_d(\tau) = q^{-d} + \sum_{n>0} A(n,d) q^n \ , \quad d \equiv 0, 3 \bmod 4 \ , \tag{5.21}$$

where $f_0 = \theta$. Here we are mainly interested in the case $d = 3$. Using uniqueness, it is easy to show that f_3 can be explicitly specified as

$$f_3(\tau) = -\frac{1}{20}\left(\frac{[\theta(\tau), E_{10}(4\tau)]}{\Delta(4\tau)} + 608\theta(\tau) \right) = q^{-3} - 248q + 26752q^4 - \dots \ , \tag{5.22}$$

where $E_{10} = E_4 E_6$ is the weight 10 Eisenstein series and Δ is the weight 12 cusp form introduced in §2.1. The *Rankin-Cohen bracket* is defined by $[f, g] := kfDg - lgDf$, with $D = \frac{1}{2\pi i}\frac{d}{d\tau}$ and k, l being the weights of the modular forms f, g, respectively.

Thompson moonshine was initiated by the observation that the first few Fourier coefficients of f_3 can be expressed in terms of the dimensions of irreducible representations of *Th*. For example $1 = 1$, $248 = 248$ and $26752 = 27000 - 248$. It is worth noting that this connection between f_3 and *Th* had already appeared in the context of generalised monstrous moonshine [74, 102, 103], cf. §4.1. Specifically, the centraliser of the class $3C$ of $\mathbb{M}$ is isomorphic to $\mathbb{Z}_3 \times Th$ [63], and consequently the twisted module $V^{\natural}_{3\mathrm{C}}$ carries a natural action of *Th*. The corresponding twisted generalised character is given by the so-called *Borcherd's lift* of f_3,

$$Z(3\mathrm{C}, e; \tau) = q^{-\frac{1}{3}} \prod_{n>0} (1 - q^n)^{A(n^2,3)} \ . \tag{5.23}$$

Notice however that the above connection to generalised monstrous moonshine is limited only to the coefficients $A(n^2, 3)$ of f_3, which does not explain why the rest of its coefficients also exhibit a relation to *Th*. To do better, the insight of [99] is to "correct" the function by including additional contributions from the theta function $\theta(\tau)$, and consider instead

$$\mathcal{F}_3(\tau) := 2f_3(\tau) + 248\theta(\tau) = 2q^{-3} + 248 + 2 \cdot 27000q^4 - 2 \cdot 85995q^5 + \dots \ , \tag{5.24}$$

which belongs to the Kohnen plus space. Note that the alternating signs correspond to the superspace structure mentioned in Conjecture 5.3.

The McKay–Thompson series $\mathcal{F}_{3,g}$ of Thompson moonshine have a structure very similar to the above. Schematically, they can be most naturally viewed as a sum of two contributions

$$\mathcal{F}_{3,g} = \text{Projected Rademacher Sum} + \theta\text{-corrections} \ , \tag{5.25}$$

where the first term is the Rademacher sum of the unique polar term $2q^{-3}$ for the suitable group Γ_g and multiplier system, projected onto the Kohnen plus space, and the second term again involves linear combinations of $\theta(d^2\tau)$, where d are certain integers compatible with the group Γ_g. We refer to [99], and in particular [104] for more details.

It is interesting to note that the modular objects in Thompson moonshine enjoy a close relation to the so-called traces of singular moduli, which are important quantities in number theory. As mentioned in the final example of §2.3, the generating functions of the traces of singular moduli turn out to have interesting modular properties [101]. In particular, we have

$$A(n,3) = \frac{1}{\sqrt{n}} \sum_{Q\in\mathcal{Q}_{-3n}/\,PSL_2(\mathbb{Z})} \chi_{n,-3}(Q)J(\alpha_Q)\ , \quad n\in\mathbb{Z}_+\ , \tag{5.26}$$

coinciding with the Fourier coefficients of the modular form f_3 which participates in Thompson moonshine. In the above, $\chi_{n,-3}$ denotes a certain genus character whose precise definition can be found in [101]. An analogous relation to traces of singular moduli can be established also at higher level, when one only keeps track of $\Gamma_0(N)$-equivalences of quadratic forms.

6 Moonshine at weight three-halves

This is an arena for moonshine that is currently in rapid development. As a result we will be very brief and just highlight a few key features and prospects of the recent developments.

After the fruitful explorations in moonshine at weight one-half, it is natural to wonder about the landscape of moonshine in the dual weight, $w = 2 - \frac{1}{2} = \frac{3}{2}$. Weight $\frac{3}{2}$ mock modular forms are much less constrained than their weight $\frac{1}{2}$ counterparts, and they turn out to provide a fertile ground to study the connections to finite groups, already giving us the *O'Nan moonshine* [105] and the *class number moonshine* [106], the latter involving the Hurwitz-Kronecker class number mock modular form described in (2.48).

It is more likely than not that there are more treasures in weight 3/2. Here we limit ourselves to commenting on the two established moonshine examples.

We will point out three conspicuous properties of the two examples of moonshine at weight three-halves[4].

Supermodules. In monstrous and Conway moonshine discussed in §4, the underlying modules are "real" modules (as opposed to supermodules (1.20)) and the group characters involved are ordinary traces (as opposed to supertraces). The same is true in Mathieu and umbral moonshine, up to the polar terms in (5.8) which have negative coefficients but do not carry any non-trivial information on group representations. In Thompson moonshine we encounter supermodules, but only in a rather trivial way: the supermodule is purely even or purely odd depending on whether the corresponding q-power is $+1$ or -1 mod 4. In particular, in weight one-half we do not encounter moonshine examples in which Fourier coefficients of mock modular forms are given by differences of traces of group representations, i.e. in every homogeneous component the supermodule is always either even or odd.

In weight three-halves this is no longer true and the super-ness of the supermodules is an intrinsic and unavoidable property of the two moonshine examples we discuss.

---

[4] In fact, these properties make the weight 3/2 connections so different from the moonshine examples at weight zero and weight one-half, such that one might wonder whether we should still use the term "moonshine" for them.

Cusp forms. In the moonshine examples we discussed so far, the (mock) modular objects involved are very distinguished in that they are essentially uniquely determined by their modularity and analyticity properties. In particular, the graded characters of monstrous and Conway moonshine are Hauptmoduls and can be written as, up to an additive constant, Rademacher sums with a single polar term q^{-1}. The graded characters of umbral (including Mathieu) and Thompson moonshine are also (projected) Rademacher sums up to addition of theta functions (cf. (5.25) and [92]). The latter can arguably be viewed as a minor correction since the coefficients of theta functions are vanishing at almost all exponents (vanishing unless the exponenets are square numbers up to a specific rescaling) and they do not carry any interesting number theoretic information.

This crucial uniqueness property, which serves to distinguish mock modular forms for weight zero and weight one-half moonshine, is no longer the relevant criterion in weight three-halves moonshine. The mock modular forms in O'Nan and class number moonshine involve weight 3/2 cusp forms in a very pronounced way. (See the end of §2.1 for the definition of cusp forms.)

This property opens the door to connecting with arithmetic geometry, a totally different area in mathematics. The main ingredients of this connection are: 1) The Waldspurger-type relation among coefficients of weight 3/2 newforms and the special L-values of the corresponding weight two newforms under the so-called Shimura correspondence [107]. 2) The modularity theorem relating the L-function of elliptic curves and that of certain weight two newforms. 3) The (strong form of) Birch and Swinnerton-Dyer conjecture (and its proven elements) relating special L-values of the elliptic curves and their important arithmetic data including the order of the Tate-Shafarevich group. See [108–110]. 4) The congruences among the moonshine modular objects stemming from the congruences among the characters which is a basic property of finite group representations. Combining the above ingredients, the connections between weight 3/2 cusp forms and finite groups lead to an infinite sequence of statements on the p-part of the Tate-Shafarevich group, organised by the finite group structure. Moreover, in class number moonshine the torsion subgroups of the elliptic curves (and other Abelian varieties in general) attached to certain modular data play an important role in the construction of moonshine mock modular forms. Exploring the implication of the finite group structure in mock modular forms in arithmetic geometry is an interesting new mathematical direction in the study of moonshine.

New finite groups. The above two properties , regarding the finite group and the modular form sides of the moonshine connection respectively, makes it possible for weight three-halves moonshine to encompass new types of finite groups into the realm of moonshine.

O'Nan moonshine, as the name suggests, connects a set of weight 3/2 mock modular forms to the O'Nan group. The O'Nan group was discovered in 1976 in the midst of the intense developments related to the classification of finite groups, and is one of the 6 pariah groups not related to the Monster. The class number moonshine, on the other hand, delivers a family of infinitely many weight 3/2 mock modular forms arising from Hurwitz–Kronecker class numbers (2.47), which accommodates representations of *infinitely many groups*. Recall that, apart from the sporadic groups there are three infinite families in the classification of finite simple groups, and one of them is the cycle groups $\mathbb{Z}_p$ of prime order p. It turns out that for each prime number p one has a mock modular form $\mathcal{H}_{p,1}$ of level 4, a scalar multiple of the class number mock modular form $\mathcal{H}$ (2.48), and a mock modular form $\mathcal{H}_{p,p}$ of level $4p$, such that they coincide with the two graded characters of a $\mathbb{Z}$-graded supermodule for $\mathbb{Z}_p$. Apart from this infinite family of groups, it has also been shown in [106] that mock modular forms at different levels (given by 4 times the order of the group elements) conspire to encode supermodules for interesting Mathieu groups.

These instances of weight three-halves moonshine greatly broaden the horizon of connections between modular objects and finite groups. All instances of moonshine at weight zero and one-half are connected to groups related to the Monster, but the O'Nan and the class number examples show that this does not have to be the case. In these examples it was demonstrated that a pariah sporadic group as well as one of the 3 infinite families of non-sporadic finite simple groups are also related to (mock) modular forms in a similar way! This suggests the tantalising possibility that moonshine-type connections are in fact ubiquitous and an intrinsic property in the landscape of modular objects and finite groups, and that the latter is an important organising structure of the former.

Last but not least, these new connections, involving supermodules of infinitely many groups, pose interesting challenges to physics. Can physics still provide the finite group modules and an explanation for these connections? If so, what are the relevant physical systems that can give rise to these supermodules?

Part III

Moonshine and string theory

String theory has been playing a key role in moonshine from its earliest years, starting with the string-inspired proof of monstrous moonshine by Borcherds (cf. §4.1). String theory is also expected to play an interesting role in the understanding of Mathieu and umbral moonshine, and we hope that moonshine can teach us something novel and important about certain aspects of string compactification. More generally speaking, the connection to physics is one of the key reasons why many string theorists and mathematicians find moonshine exciting. In this part we will discuss some of these connections. In §7 we outline the relation between monstrous moonshine and certain (1+0)-dimensional string compactification. In §8, the main section of this part, we summarise the known relations between various moonshines and $K3$ compactifications. Finally we briefly mention some other connections discussed recently in §9.

7 Monstrous moonshine

In this section we briefly describe a recent string theory construction, developed in [111, 112], involving the monstrous moonshine module $V^\natural$ and the associated twisted module $V_g^\natural$ discussed in §4.1. Although the striking genus zero property of monstrous moonshine has been mathematically proven, its physical interpretation has remained elusive for a long time. The recent work [111, 112] sheds light on this important property by considering certain heterotic string theory compactifications down to $0 + 1$ dimensions.

In [111, 112], one takes as the starting point a compactification down to $1 + 1$ dimensions where the internal CFT, which has central charge $(24, 12)$, is given by $V^\natural \times V^{s\natural}$. The second factor is the Conway module introduced in §4.2, which plays mostly the role of the spectator whose main function is to provide the right-moving $\mathcal{N} = 1$ supersymmetry of the world-sheet theory. This model turns out to have $\mathcal{N} = (0, 24)$ spacetime supersymmetry in two-dimensions. We further compactify the remaining spatial direction on a circle. This additional compactification is crucial for computing the "BPS index" described below and for the construction of more general models. More generally, for each $g \in \mathbb{M}$ one constructs a CHL-like model by orbifolding the theory by a symmetry which acts on $V^\natural$ as g and on the circle as a shift of order dictated by certain properties of the action of g on $V^\natural$ (and coincides with the order of g in the simplest cases). To avoid subtleties involving gravitational anomalies, this theory should be considered at the zero-string coupling limit.

In [111] it is shown that the (first-quantized) BPS states of the above models form a module of the generalised Kac-Moody algebras $\mathfrak{m}_\mathfrak{g}$, which is the monster algebra constructed by Borcherds for $g = e$ and the "twisted" monster algebras constructed by Carnahan for $g \neq e$ [75, 103, 113]. See also the discussion on generalised monstrous moonshine in §4.1. We are interested in the (twined) BPS index counting spacetime BPS states in these theories. Denote the relevant Hilbert space by $\mathcal{H}^g_{\text{BPS}}$, we define

$$Z_{(g,h)}(T, U) := \text{Tr}_{\mathcal{H}^g_{\text{BPS}}} \left((-1)^F e^{2\pi i TW} e^{2\pi i UM}\right) \ , \tag{7.1}$$

where g, h are commuting elements of the Monster. In the above, $T, U \in \mathbb{H}$ are complexified chemical potentials for the winding and momentum quantum numbers W, M respectively, with the imaginary parts scaling according to the inverse temperature and radius as

$$\mathrm{Im}T \sim \beta R\ , \quad \mathrm{Im}U \sim \frac{\beta}{R}\ . \tag{7.2}$$

Using the Fock space construction of $\mathcal{H}^g_{\mathrm{BPS}}$ by allowing for an arbitrary number of free strings, as well as the twisted denominator formula, relating an infinite product to an infinite sum and generalising the Koike–Norton–Zagier formula (4.14), one obtains

$$\begin{aligned} Z_{(e,g)}(T,U) &= \left(T_{(e,g)}(T) - T_{(e,g)}(U)\right)^{24} \\ Z_{(g,e)}(T,U) &= \left(T_{(e,g)}(T) - T_{(g,e)}(U)\right)^{24}\ . \end{aligned} \tag{7.3}$$

Note that $T_{(e,g)} = T_g$ is the McKay–Thompson series of the monstrous moonshine.

The above expressions lead to a physical interpretation of the modularity and the genus zero property of monstrous moonshine. First, it was shown in [111], using the techniques developed in [114], that the self-duality group (i.e., a subgroup of the T-duality group mapping between the same CHL model but generically at different values of the moduli) involves two subgroups of $SL_2(\mathbb{R})$, acting on T and U independently in (7.3). In particular, the $SL_2(\mathbb{R})$ subgroup acting on T is nothing but Γ_g, the associated genus zero group in monstrous moonshine (cf. Conjecture 4.1). As a result, the physical invariance of the BPS index under self-duality group immediately leads to the Γ_g-invariance modularity property of the moonshine function $T_g(T)$. Second, recall that the Hauptmodul property of $T_{(e,g)}$ can be recast in the claim that they have only a single pole at $\tau \to i\infty$ and its images under Γ_g. In terms of the supersymmetric index, this means that $Z_{(g,e)}(T,U)^{1/24}$ and its T-duality images will diverge as $T \to i\infty$, for fixed U. Using the T-duality

$$Z_{(e,g)}(T,U) = -Z_{(e,g)}\left(-\frac{1}{U}, -\frac{1}{T}\right)$$

and the generalised moonshine relation between $T_{(g^ah^c,g^bh^d)}(\tau)$ and $T_{(g,h)}\left(\frac{a\tau+b}{c\tau+d}\right)$ (cf. §4.1), this property gets translated to the statement that for all co-prime integers a and b, $Z_{(g^a,g^b)}(T,U)$ is divergent in the limit $U \to i\infty$ with T fixed, if and only if this limit is related to the limit of $T \to i\infty$ with U fixed by a self-duality of the CHL model associated to $g \in \mathbb{M}$. Now we will argue why this is true from physical considerations. Let us interpolate between the two limits by varying the radius R, while fixing the inverse temperature β to be large. At the low temperature limit, the divergence at the limit $T \to i\infty$ with U fixed is caused by the ground state contribution $e^{\beta R}$, which blows up as $R \to \infty$. The small radius limit is convergent unless there is a phase transition and a new ground state with contribution $e^{\beta/R}$ emerges at the other side of the critical line. If this happens, one can show that there is an $SU(2)$ symmetry at the self-dual radius limit on the critical line, which in particular contains the T-duality. In other words, the two sides of the critical line are related by a self-duality contained in Γ_g, and this proves the Hauptmodul property.

8 Moonshine and $K3$ string theory

As mentioned in §5, Mathieu moonshine was first noticed in the context of $K3$ sigma models and was later recognised as being one of the twenty-three cases of umbral moonshine. As we will discuss shortly, all instances of umbral moonshine, not just Mathieu moonshine, enjoy a close relation to $K3$ elliptic genus, the quantity that captures the BPS states of $K3$ sigma models, and more generally BPS states of superstring theory compactifications involving $K3$ surfaces. This tentative connection to $K3$ string theories is what makes Mathieu and umbral moonshine so interesting for many string theorists.

This is because $K3$ manifolds have the special property that they are the only non-trivial (i.e. non-toroidal) Calabi–Yau manifolds in two (complex) dimensions, and the resulting string theory compactifications are much more manageable than compactifications on generic Calabi–Yau three-folds due to higher amount of supersymmetries, and contain much richer structure and information than toroidal compactifications which are for many purposes too simple to be useful. As a result, $K3$ compactifications have been playing a prominent role throughout the development of string theory, being the playground for developing revolutionary new ideas including black hole microstates counting [115] and AdS/CFT correspondence [116]. Furthermore, they also feature in important dualities relating different string theories.

In this section we will review the observed relation between moonshine and $K3$ sigma models, and outline the challenges to be overcome in order to understand the origin of this relation and to find a physical system which provides the much wanted uniform construction of the umbral moonshine module. Finally, we will summarise the different ideas that aim to address these challenges, along with their current status.

8.1 Mathieu moonshine and $K3$ elliptic genus

As mentioned in §3.3, the Calabi-Yau property of $K3$ together with the Jacobi form property dictates that its elliptic genus is given by $2\phi_{0,1}$ in terms of the familiar weak Jacobi form (2.31). Recall that $K3$ is the only Calabi-Yau two-fold besides the four-torus. It possesses the special property that it is not only Kähler but also admits a hyper-Kähler structure. As a result, the $U(1)$ symmetry of the $\mathcal{N}=2$ superconformal algebra discussed in §3.3 can be extended to $SU(2)$, and the sigma model has enhanced $\mathcal{N}=4$ superconformal symmetry for both the left and the right movers. This leads to a specific decomposition of the elliptic genus of an $\mathcal{N}=4$ SCFT that we will now explain.

Since the underlying CFT admits an action by the $\mathcal{N}=4$ superconformal algebra, the Hilbert space decomposes into a direct sum of (unitary) irreducible representations of this algebra. Hence the elliptic genus (3.34) can be written as a sum of characters of representations of this algebra with some multiplicities. A natural embedding of the $U(1)$ current algebra of the $\mathcal{N}=2$ superconformal algebra into the $SU(2)$ current algebra of the $\mathcal{N}=4$ superconformal algebra is obtained by choosing $J_0^3 \sim J_0$. As a result, the $\mathcal{N}=4$ highest weight representations are again labelled by two quantum numbers h,ℓ, corresponding to the operators L_0, J_0^3 respectively, and the character of such an irreducible representation $V_{h,\ell}$ is defined as

$$ch_{h,\ell}(\tau,z) := \mathrm{Tr}_{V_{h,\ell}}\left((-1)^{J_0}y^{J_0}q^{L_0-c/24}\right).$$

For the central charge $c=6$, there are two supersymmetric (also called ‘BPS’ or ‘massless’) representations in the Ramond sector, and they have the quantum numbers

$$h = \tfrac{1}{4}\ ,\ \ell = 0, \tfrac{1}{2}\ .$$

Their characters have been computed to be [117–119]

$$\begin{aligned} ch_{\frac{1}{4},0}(\tau,z) &= \frac{\theta_1^2(\tau,z)}{\eta^3(\tau)}\mu(\tau,z)\,, \\ ch_{\frac{1}{4},\frac{1}{2}}(\tau,z) &= q^{-\frac{1}{8}}\frac{\theta_1(\tau,z)}{\eta^3(\tau)} - 2\frac{\theta_1(\tau,z)^2}{\eta^3(\tau)}\mu(\tau,z)\,, \end{aligned} \tag{8.1}$$

where $\mu(\tau,z)$ is the so-called Appell-Lerch sum defined in (5.2).

Notice that the supersymmetric representations have non-vanishing Witten index

$$ch_{\frac{1}{4},0}(\tau,z=0) = 1\,, \quad ch_{\frac{1}{4},0}(\tau,z=0) = -2\,.$$

The non-supersymmetric (or ‘non-BPS’ or ‘’) representations have

$$h = \tfrac{1}{4} + n\,, \quad \ell = \tfrac{1}{2}\,, \qquad n = 1,2,\dots,$$

and their characters are given by

$$ch_{\frac{1}{4}+n,\frac{1}{2}}(\tau,z) = q^{-\frac{1}{8}+n}\frac{\theta_1^2(\tau,z)}{\eta^3(\tau)}\,. \tag{8.2}$$

By definition, their contribution to the Witten index vanishes, $ch_{\frac{1}{4}+n,\frac{1}{2}}(\tau,z=0) = 0$. This is why the Witten index only receives contribution from quantum states that are of the form ($\text{massless}_L \times \text{massless}_R$) in terms of $\mathcal{N}=4$ representations, while the elliptic genus receives contribution from ($\text{massless}_L \times \text{massless}_R$) as well as ($\text{massive}_L \times \text{massless}_R$).

We are now ready to describe the first physical interpretation of the specific way (5.1) of decomposing the $K3$ elliptic genus, which makes the connection to M_{24} manifest. From (5.1) and (8.1)-(8.2), we see that $\mathbf{EG}(K3)$ can be expressed as

$$\mathbf{EG}(K3) = 20\,ch_{\frac{1}{4},0} - 2\,ch_{\frac{1}{4},\frac{1}{2}} + \sum_{n=1}^{\infty} c_H(n)\,ch_{\frac{1}{4}+n,\frac{1}{2}} \tag{8.3}$$

where

$$H(\tau) = q^{-1/8}\left(-2 + \sum_{n=1}^{\infty} c_H(n)q^n\right) = 2q^{-\frac{1}{8}}\left(-1 + 45\,q + 231\,q^2 + 770\,q^3\dots\right)\ , \tag{8.4}$$

is the mock modular form given in (2.42). In particular, the multiplicities $c_H(n)$ of the massive multiplets are equal to the dimensions of the homogeneous component of the M_{24}-module K according to Mathieu moonshine Conjecture 5.2.

As mentioned already in §5.1, it is then tempting to identify ϕ_g (5.3), encoding the McKay–Thompson series of Mathieu moonshine H_g, as the elliptic genus twined by certain symmetries of the $K3$ non-linear sigma model whose action preserves the $\mathcal{N}=4$ SCA. While this is possible for some $g \in M_{24}$, this interpretation is proven to fail for some other group elements. We will discuss this in §8.4.

8.2 Umbral moonshine and $K3$ elliptic genus

As alluded to before, the $\mathcal{N}=4$ decomposition is not the only possible interpretation for the splitting (5.1) of the elliptic genus, into the part coming directly from the moonshine mock modular form and the contribution from the Appell-Lerch sum μ. A second interpretation, in terms of the elliptic genus of singular spaces labelled by ADE root systems, yields the same result in this case but the difference is that it can be readily generalised to the other 22 cases of umbral moonshine.

Recall that, for each of the 23 Niemeier lattices N^X, uniquely determined by its root system X, umbral moonshine associates a finite group G^X and a weight one mock Jacobi form $\psi_g^X = \sum_r H_{g,r}^X \hat{\theta}_{m,r}$ for each $g \in G^X$. In [120], it was shown that there are 23 ways of splitting $\mathbf{EG}(K3)$ into two part:

$$\mathbf{EG}(K3;\tau,z) = \mathbf{EG}(X;\tau,z) + \frac{\theta_1^2(\tau,z)}{2\eta^6(\tau)}\left(\frac{1}{2\pi i}\frac{\partial}{\partial w}\psi^X(\tau,w)\right)\Big|_{w=0} \tag{8.5}$$

where $\psi^X = \psi_e^X$ is the optimal mock Jacobi form given in §5.2, encoding the graded dimension of the umbral moonshine module. In the above formula, $\mathbf{EG}(X;\tau,z)$ denotes the holomorphic part of the elliptic genus of the singularities corresponding to the root system of N^X. See also the remark at the end of §3.3. Note that a common definition of the $K3$ surface is that it has at worst du Val type surface singularities, i.e. singularities of the complex plane of the form $\mathbb{C}^2/G$, where G is a finite subgroup of $SU(2)_{\mathbb{C}}$. These singularities have an ADE classification, in accordance with the famous McKay correspence. A conformal field theory description of string theory with these ADE singularities as the target space was given in [121]. The form of their elliptic genus was investigated in a number of papers, including [24,27,29–31,120,122]. For instance, when N is the Niemeier lattice with root system $24A_1$, we have

$$\mathbf{EG}(N;\tau,z) := 24\,\mathbf{EG}(A_1;\tau,z)\ ,$$

defined to be 24 times the holomorphic part of the elliptic genus of an A_1-singularity. One has

$$\mathbf{EG}(A_1;\tau,z) = ch_{\frac{1}{4},0}(\tau,z) = \frac{\theta_1^2(\tau,z)}{\eta^3(\tau)}\mu(\tau,z). \tag{8.6}$$

As a result, the splitting (5.1) of the elliptic genus can also be interpreted as coming from two contributions, one from the singularity elliptic genus corresponding to the root system $X = 24A_1$, and one from the umbral moonshine function ψ^{24A_1}. Different from the $\mathcal{N}=4$ character interpretation, this interpretation of the splitting is applicable for all 23 cases of umbral moonshine. In other words, (8.5) is really 23 equalities, corresponding to all 23 Niemeier lattices N^X.

This interpretation suggests a twined version of the function $\mathbf{EG}(K3;\tau,z)$ by twining the right-hand side of the equality (8.5):

$$\phi_g^X(\tau,z) := \mathbf{EG}_g(X;\tau,z) + \frac{\theta_1^2(\tau,z)}{\eta^6(\tau)}\left(\frac{1}{2\pi i}\frac{\partial}{\partial w}\psi_g^X(\tau,w)\right)\Big|_{w=0}\ , \tag{8.7}$$

where g can be any element of G^X. The first term on the right-hand side is the (holomorphic part of the) twined elliptic genus of the corresponding singularity sigma model. This can be computed explicitly, since we know the G^X–action on the Niemeier root system X, which translates into an G^X–action on the singularity

CFT that preserves its superconformal structure, and leads to a definition of its twined elliptic genus. The second term, on the other hand, is directly provided by the McKay–Thompson series of umbral moonshine. It is easy to check that this definition of ϕ_g^X for $N = N^{24A_1}$ leads to the same weak Jacobi forms as in (5.3) for the case of Mathieu moonshine. Just like in the case of Mathieu moonshine, it is then tempting to interpret all the functions ϕ_g^X arising from umbral moonshine as in (8.7), as the elliptic genus twined by certain symmetries of the $K3$ non-linear sigma model whose action preserves the $\mathcal{N} = 4$ SCA.

For some of the $g \in G^X$, such an interpretation is not availble, as we will discuss in §8.4. In particular, only subgroups of G^X which preserve a four-plane, i.e. a four-dimensional oriented subspace within $N^X \otimes_{\mathbb{Z}} \mathbb{R}$, will acquire a role as a symmetry group of $K3$ sigma models. For later use, for a given Niemeier lattice $N = N^X$ with non-trivial root system, we will denote the set of Jacobi forms arising from (8.7) by

$$\Phi(N^X) := \{\phi_g^X \mid g \text{ is a four-plane preserving element of } G^X\}\ . \tag{8.8}$$

8.3 Conway moonshine and $K3$ elliptic genus

Recall the construction of the twisted Conway module $V_{\rm tw}^{s\natural}$ described in §4.2. Given a fixed n-dimensional subspace in $\Lambda \otimes_{\mathbb{Z}} \mathbb{C}$, there are different ways to build $U(1)$ currents by considering bilinear combinations of the 24 fermions [78, 79]. In this theory, one can construct a $U(1)$ current J of level 2 from fermions associated to a four-plane, i.e. a subspace of (real) dimension four. Given such a four-plane Π, fixing the $U(1)$ current and the SVOA structure, breaks the symmetry of $V_{\rm tw}^{s\natural}$ from Co_0 to its subgroup preserving Π. Conversely, given a four-plane preserving $G < Co_0$, one can construct a $U(1)$ current J such that $V_{\rm tw}^{s\natural}$, when equipped with a module structure for J and for the $\mathcal{N} = 1$ superconformal algebra, has symmetry G.

Interestingly, the $U(1)$-charged graded character of $V_{\rm tw}^{s\natural}$ coincides with $\mathbf{EG}(K3)$ (up to a sign) [78]. More generally, one can consider the $U(1)$-graded character of the twisted Conway module twined by a four-plane preserving element of Co_0 (see §4.2 for notation):

$$\phi_g^\Lambda := -\,\mathrm{Tr}_{V_{\rm tw}^{s\natural}}\left[\mathfrak{z}\ \hat{g}\ y^{J_0} q^{L_0-\frac{c}{24}}\right], \tag{8.9}$$

where J_0 is the zero mode of the $U(1)$ current. Explicitly, it is given by

$$\begin{aligned}\phi_g^\Lambda(\tau,\zeta) = \frac{1}{2}\Bigg[&\frac{\theta_3(\tau,\zeta)^2}{\theta_3(\tau,0)^2}\frac{\eta_{-g}(\tau/2)}{\eta_{-g}(\tau)} - \frac{\theta_4(\tau,\zeta)^2}{\theta_4(\tau,0)^2}\frac{\eta_g(\tau/2)}{\eta_g(\tau)} \\ &- \frac{\theta_2(\tau,\zeta)^2}{\theta_2(\tau,0)^2} C_{-g}\eta_{-g}(\tau) - \frac{\theta_1(\tau,\zeta)^2}{\eta(\tau)^6} D_g \eta_g(\tau)\Bigg]\ ,\end{aligned} \tag{8.10}$$

where most special functions were defined in (4.28). The function D_g is defined by

$$D_g := \nu \sideset{}{'}\prod_{i=1}^{12}\left(1-\lambda_i^{-1}\right) = \sideset{}{'}\prod_{i=1}^{12}\left(\nu_i - \nu_i^{-1}\right). \tag{8.11}$$

where in $\prod'$ one skips the two pairs of eigenvalues associated with the fixed four-plane, for which $\lambda_i^{\pm 1} = 1$. Notice that D_g is non-vanishing if and only if it fixes exactly a four-plane and not more. In the latter case, D_g is determined up to a sign by the eigenvalues of g, since we are free to exchange what we call λ_i and λ_i^{-1}. As a result, for exactly four-plane preserving elements there are in fact two choices of ϕ_g^Λ, depending on the choice of the sign of D_g, and we define

$$\begin{aligned}\phi_{g,\epsilon}^\Lambda(\tau,\zeta) := \frac{1}{2}\Bigg[&\frac{\theta_3(\tau,\zeta)^2}{\theta_3(\tau,0)^2}\frac{\eta_{-g}(\tau/2)}{\eta_{-g}(\tau)} - \frac{\theta_4(\tau,\zeta)^2}{\theta_4(\tau,0)^2}\frac{\eta_g(\tau/2)}{\eta_g(\tau)} \\ &-\frac{\theta_2(\tau,\zeta)^2}{\theta_2(\tau,0)^2}C_{-g}\eta_{-g}(\tau) - \epsilon\,\frac{\theta_1(\tau,\zeta)^2}{\eta(\tau)^6}\,|D_g|\,\eta_g(\tau)\Bigg]\ ,\end{aligned} \tag{8.12}$$

where $\epsilon = \pm 1$ encodes the sign ambiguity of D_g.

In all cases, it was shown in [78] that the $\phi_{g,\epsilon}^\Lambda$ are Jacobi forms of weight 0 and index 1, at some level, for every $g \in Co_0$ that fixes at least a four-plane. For future use, denote the set of all these functions by

$$\Phi(\Lambda) := \left\{\phi_{g,+}^\Lambda, \phi_{g,-}^\Lambda \mid g \text{ is a four-plane preserving element of } Co_0\right\}\ . \tag{8.13}$$

We will mention a few observations about this set of functions and those arising from umbral moonshine, the $\Phi(N^X)$ defined in (8.8). First, these sets of functions notably have a large overlap: quite often Conway and umbral moonshine give rise to the same weak Jacobi form, albeit via very different ways. An important difference is that a weak Jacobi form ϕ_g^X can be defined for all $g \in G^X$ via umbral moonshine (8.7), while in the Conway case one has to restrict to four-plane preserving elements (in order to obtain a Jacobi form of index one). Second, the functions that only appear in $\Phi(\Lambda)$ tend to be realised as twined elliptic genera of $K3$ sigma models on the sub-loci in the moduli space where the model admits a description as T^4-orbifold. Note also that the geometry of T^4-orbifold connects nicely to the free field orbifold construction of the Conway module. In fact, $V_{\rm tw}^{s\natural}$ is isomorphic to the $\mathbb{Z}_2$ orbifold of a certain T^4 model, when one "forgets" about the distinction between left- and right-movers, roughly speaking [78, 123, 124]. Third, the functions that only arise from umbral and not from Conway moonshine have the property that they must correspond to a symmetry which acts differently on the left- and right-movers of a $K3$ sigma model [125]. We will discuss the connections between the moonshine symmetry groups and the stringy $K3$ symmetry groups in the next subsection.

8.4 Symmetries of $K3$ string theory and moonshine

In the previous subsections we have seen a close mathematical relation between the $K3$ elliptic genus and the functions of umbral and Conway moonshine. It is therefore natural to compare the moonshine groups and the symmetries of $K3$ sigma models. In fact, it was known that symmetries of $K3$ surfaces have a close relation to sporadic groups. In particular, a celebrated theorem by Mukai [126] established the role of M_{23}, one of the 26 sporadic groups and a subgroup of M_{24}, as the finite group that "organises" the symmetries of $K3$ surfaces. More precisely, in [126] it was shown that there is of a bijection between (isomorphism classes of) M_{23} subgroups with at least five orbits and (isomorphism classes of) finite groups of $K3$ symplectomorphisms. A generalisation of this classical result to "stringy $K3$ geometry" was initiated by Gaberdiel, Hohenegger, and Volpato, who in [127] classified the symmetry groups of all (non-singular) $K3$ sigma model using lattice techniques in a method closely following the simplified proof of the Mukai theorem by Kondo [128]. From the spacetime (D-branes) point of view, the results of [127] can be viewed as classifying symplectic autoequivalences (symmetries) of derived categories on $K3$ surfaces [129]. See also [130] for a related discussion on symmetries of appropriately defined moduli spaces relevant for curve counting on $K3$. In [125] the classification was extended to theories corresponding to singular loci in the moduli space of $K3$ sigma models. This is necessary if one is interested in the full string theory and not just the sigma models. Despite the fact that the type IIA worldsheet theory behaves badly along these loci [131], the full type IIA string theory is not only completely well-defined but also possesses special physical relevance in connection to non-Abelian gauge symmetries. The spacetime gauge group is enhanced from $U(1)^{24}$ to some non-Abelian group at these loci, and the ADE type gauge group is given by the ADE type singularity of the $K3$ surface [131,132]. The existence of such loci with enhanced gauge symmetries in the moduli space, though not immediately manifest from the world-sheet analysis in type IIA, is clear from the the dual description in terms of heterotic T^4 compactification.

To state the classification result, let us first review some general properties of sigma models on K3 (see [133,134]). The moduli space of non-singular non-linear sigma models on $K3$ with $\mathcal{N} = (4,4)$ supersymmetry is given by an open subset in

$$\mathcal{M} = (SO(4) \times O(20)) \backslash O^+(4,20)/O^+(\Gamma^{4,20}) \; , \tag{8.14}$$

where $(SO(4) \times O(20)) \backslash O^+(4,20)$ is the Grassmannian of positive four-planes Π, a four-dimensional oriented positive-definite subspace within $\mathbb{R}^{4,20} \cong \Gamma^{4,20} \otimes_{\mathbb{Z}} \mathbb{R}$, and $\Gamma^{4,20}$ is the even unimodular lattice with signature $(4,20)$. In other words, choosing a point in the moduli space $\mathcal{M}$ is equivalent to choosing a four-plane inside $\Gamma^{4,20} \otimes_{\mathbb{Z}} \mathbb{R}$. The complement in $\mathcal{M}$ of this open subset is the set of singular four-planes Π (i.e. orthogonal to a root $v \in \Gamma^{4,20}$, $v^2 = -2$) and correspond to certain singular limits of sigma models on K3. The whole space $\mathcal{M}$ is also the moduli space of type IIA string theory at a fixed finite value of g_s. In the singular cases, the full supersymmetry-preserving symmetries of the corresponding type IIA string theory contains a continuous non-Abelian gauge symmetry which we will mod out to obtain the discrete symmetry group G_{IIA}. As a result, we have

$$G_{\text{IIA}}(\Pi) = \text{Stab}(\Pi)/\text{W} \tag{8.15}$$

where Stab(Π) is the subgroup of $O^+(\Gamma^{4,20})$ which stabilises the four-plane Π point-wise, and W is the Weyl group generated by reflections with respect to all root vectors orthogonal to Π, if there are any.

In [125] (Corollary 4) it was then shown that the groups G_{IIA}, realised somewhere in $\mathcal{M}$, are in bijection with the four-plane preserving subgroups of the twenty-four finite groups G_N associated to the twenty-four even self-dual lattices of rank twenty-four (cf. (1.13)), including the twenty-three umbral groups G^X and the Conway group Co_0. In particular, it means that generically the umbral group G^X, and in particular $G^{24A_1} \cong M_{24}$, cannot be the symmetry group of any individual $K3$ sigma model as they are not four-plane preserving subgroups of Co_0.

After classifying the symmetry groups as abstract groups, we would like to know how they act on the spectrum of supersymmetric states. This information is captured by the elliptic genera twined by these symmetries, as defined in (3.35). In [125], it was conjectured (among other finer conditions) that any twined $K3$ elliptic genus realised anywhere in the moduli space $\mathcal{M}$ coincides with at least one of the functions in one of the twenty-four sets of functions $\Phi(N)$ defined in (8.8) and in (8.13), where N is either one of the twenty-three Niemeier lattices N^X or the Leech lattice Λ. Conversely, it was conjectured that *all* elements of $\Phi(N)$ play a role in capturing the symmetries of BPS states of $K3$. Strong evidence for the above conjecture was recently found in [135] by considering string theories closely related (by orbifolding) to type IIA string theory compactified on $K3 \times T^2$, and requiring that the analyticity structure of the BPS-counting functions is compatible with the physics of wall-crossing. It is however not yet understood why umbral moonshine functions should have such a close relationship to $K3$ sigma model twined elliptic genera. See §8.5 for more discussions.

8.5 Beyond perturbative string theory

The most important lesson from the previous subsection is that supersymmetry-preserving symmetries of $K3$ CFTs are closely related to the symmetries of umbral moonshine, but fall short of explaining umbral moonshine: many umbral groups simply cannot be realised as symmetries of $K3$ sigma model anywhere in the moduli space. In particular, while the conjectures discussed in §8.4 state that all umbral and Conway moonshine functions associated to four-plane preserving group elements play a role as the twining genera of $K3$ sigma models, and in fact capture them completely, the physical relevance of the umbral (including Mathieu) moonshine functions corresponding to group elements preserving only a two-plane remains unclear. There is at the moment no consensus regarding what the precise physical setup relevant for umbral moonshine is. At the same time, it is precisely this necessity to go beyond perturbative string theory, and the possibility of having moonshine involved in non-perturbative setups such as stringy black holes, that makes the possible connections between umbral moonshine and string theory so exciting. Here we briefly discuss a few possibilities that have been discussed to different degrees in the literature.

Combining symmetries. The idea is to find a way to combine symmetries realised at different points in the moduli space and in this way generate a larger group which also contains two-plane preserving elements. This possibility was first raised as a question "Is it possible that these automorphism groups at isolated points in the moduli space of $K3$ surface are enhanced to M_{24} over the whole of moduli space when we consider the elliptic genus?" in [81]. It is motivated by the fact that the elliptic genus receives only contributions from BPS states and is invariant across the moduli space. Concrete steps towards realising this idea in the context of Kummer surfaces, i.e. $K3$ surfaces obtained as torus orbifolds $T^4/\mathbb{Z}_2$, were taken in [136–138] and further explored in [139]. In particular, it was shown that it is possible to combine the geometric symmetries of three specific Kummer surfaces, given by the groups

$$G_0 := \mathbb{Z}_2^4 \rtimes (\mathbb{Z}_2 \times \mathbb{Z}_2) \ , \quad G_1 := \mathbb{Z}_2^4 \rtimes A_4 \ , \quad G_2 := \mathbb{Z}_2^4 \rtimes S_3 \ . \tag{8.16}$$

into the so-called octad group $G = \mathbb{Z}_2^4 \rtimes A_8$. This is a maximal subgroup of M_{24} but too large to be the automorphism group of a single $K3$ sigma model. The work of [139] supplied evidence that G indeed acts on the twisted sector BPS states that survive after moving away from the Kummer loci in the moduli space. This suggests the possibility that G, a maximal subgroup of M_{24}, indeed acts as an "overarching" symmetry within the Kummer moduli space. Finally, see also [140] for a exploration of this idea in the context of Landau–Ginzburg models. A challenging aspect of this approach is that it is technically very hard to follow a path in the moduli space to a generic point which does not correspond to a torus orbifold or admits a

Landau–Ginzburg UV description. As a result it is not straightforward to combine symmetries which are realised in a generalised $K3$ sigma model.

Lower-dimensional compactifications. A second approach is to consider string compactifications where larger groups are realised at given points in the moduli space as symmetry groups of the full string theory (not just for the perturbative string states and not just the BPS sector). For theories with 16 supercharges, this is only possible for compactifications with less than six non-compact dimensions. This idea was first discussed in [83], where the implications of Mathieu moonshine for four-dimensional BPS black holes, arising in type II superstring theory compactified on $K3 \times T^2$, were explored. This idea was futher developed in [141], and has led to nice new insights into string dualities [114].

More recently, it was shown that there are points in the moduli space of string theory compactifications to three dimensions which admit the Niemeier groups as discrete symmetry groups [142]. In the type IIA frame, these are given by compactifications on $K3 \times T^3$. The action of these symmetry groups on the 1/2-BPS states of the theory has been analyzed [142], and it would be interesting to understand the action on the 1/4-BPS states. However, it is currently unclear how functions related to the umbral moonshine McKay–Thompson series can be obtained in this physical setup. Recently, a further compactification on $K3 \times T^4$ was also explored in [143].

Heterotic $K3$ compactifications. The plausibility of the relevance of this setup lies in the fact that its moduli space contains the moduli space of $\mathcal{N} = (4,4)$ $K3$ CFT as a sub-locus. It is hence imaginable that moving away from the sub-loci will allow us to see more symmetries. At a technical level, we have less means to compute the spectrum due to smaller number of supersymmetries. This route has been somewhat explored in [144, 145], but a lot remains to be done.

Five-brane dynamics. The connection between umbral moonshine and double scaled little string theory, describing the perturbative dynamics in the presence of NS five-branes in type IIB superstring theory, has been investigated in [146, 147]. This idea is natural in that it incorporates new non-perturbative elements (NS fivebranes) into the physical setup, and that the double scaled little string theories admit an ADE classification which also plays an important role in the construction of umbral moonshine. It would be interesting to understand the finite group symmetries present in this setup.

Relations to certain VOA. As we have seen in the previous subsections, umbral and Conway moonshine are related, as they are both closely related to $K3$ sigma models. In [79], a close variant of the Conway module is shown to exhibit an action of a variety of two-plane preserving subgroups of Co_0, including M_{23}, and yields as twining genera a set of weak Jacobi forms of weight zero and index two. In addition, the mock modular forms which display M_{23} representations appear to be very closely related to the mock modular forms which play a role in M_{24} moonshine. It might be fruitful to gain a deeper physical understanding of this apparent connection and see if one can "tweak" the better understood Conway module in some ways to accommodate the umbral modules.

Total unknown. As none of the above approaches has led to a definite answer so far, a possible scenario is that the connections we observed between $K3$ elliptic genera and umbral moonshine are just a coincidence and the physical context (if any) of umbral moonshine lies completely somewhere else.

Finding the correct physical setup in which the umbral moonshine modules can be constructed in a natural and uniform way is an active direction area and is currently the holy grail in the study of umbral moonshine.

9 Other connections

MSW strings. Another physical setup which delivers examples of modular objects connected to finite groups is M5 branes wrapping divisors of Calabi–Yau threefolds [18]. These give rise to effective strings, the so-called MSW strings, with (0,4) worldsheet supersymmetry [19]. In [18], it was pointed out that the generalised genera of the MSW string theories are examples of skew-holomorphic Jacobi forms (cf. §2.2), a type of modular object playing a starring role in the family of moonshine generalising the Thompson moonshine, but whose role in physics has thus far not been highlighted. In particular, the generalised elliptic genera for one or two M5 branes wrapping the surfaces $\mathbb{P}^2$, del Pezzo 8, and half-$K3$ were examined in [18]. In the first case, it was known that two M5 branes wrapping the surface $\mathbb{P}^2$ leads to mock skew-holomorphic forms that are closely related to the generating function of the class number mock modular form (2.48). This connection suggests the possibility that M5-branes on $\mathbb{P}^2$ give a starting point from which we may pursue a geometric understanding of Mathieu moonshine for (rescaled) Hurwitz class numbers described in §6.

For the other cases, one finds that a weight 3/2 modular form governing the generalised elliptic genus of a single M5 brane wrapping the del Pezzo 8 and half-$K3$ surfaces is $f^{(1)} = E_4\eta^{-5}$. See §2.1 for the definition of E_4 and η. In [18] it was shown that $f^{(1)}$ is the graded superdimension of a supermodule for the Mathieu group M_{12}, and the corresponding graded characters $f_g^{(1)}$ for all $g \in M_{12}$ are explicitly given. It would be very interesting to have a geometric and string theoretic understanding of this M_{12}-supermodule.

References

[1] M. Hamermesh, *Group theory and its applications to physical problems / by M. Hamermesh.* Addison-Wesley series in physics, Reading, MA [etc.: Addison-Wesley, 1962.

[2] H. F. H. F. Jones, *Groups, representations, and physics / H.F. Jones.* Bristol; Philadelphia: Institute of Physics Pub., 2nd ed. ed., 1998.

[3] P. Webb, *A course in finite group eepresentation theory.* Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2016.

[4] P. Etingof, O. Golberg, S. Hensel, T. Liu, A. Schwendner, D. Vaintrob, and E. Yudovina, "Introduction to representation theory," *ArXiv e-prints*, Jan. 2009.

[5] D. Gorenstein, *Finite simple groups : an introduction to their classification / Daniel Gorenstein.* The university series in mathematics, New York London: Plenum, 1982.

[6] R. Solomon, "A brief history of the classification of the finite simple groups," *Bulletin of the American Mathematical Society*, vol. 38, no. 3, pp. 315–352, 2001.

[7] H.-V. Niemeier, "Definite quadratische Formen der Dimension 24 und Diskriminante 1," *J. Number Theory*, vol. 5, pp. 142–178, 1973.

[8] J. Leech, "Notes on sphere packings," *Canad. J. Math.*, vol. 19, pp. 251–267, 1967.

[9] J. H. Conway, "A characterisation of Leech's lattice," *Invent. Math.*, vol. 7, pp. 137–142, 1969.

[10] J. Leech, "Some sphere packings in higher space," *Canad. J. Math.*, vol. 16, pp. 657–682, 1964.

[11] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups.* Springer-Verlag, 1999.

[12] G. Mason, "Frame-shapes and rational characters of finite groups," *Journal of Algebra*, vol. 89, no. 2, pp. 237–246, 1984.

[13] *The 1-2-3 of modular forms : lectures at a summer school in Nordfjordeid, Norway / Jan Hendrik Bruinier ... [et al.].* Universitext, Berlin [etc.: Springer, 2008.

[14] R. E. Borcherds, "Reflection groups of Lorentzian lattices," *Duke Math. J.*, vol. 104, no. 2, pp. 319–366, 2000.

[15] M. C. Cheng and J. F. Duncan, "Optimal Mock Jacobi Theta Functions,"

[16] A. Dabholkar, S. Murthy, and D. Zagier, "Quantum Black Holes, Wall Crossing, and Mock Modular Forms," 2012.

[17] M. Eichler and D. Zagier, *The theory of Jacobi forms.* Birkhäuser, 1985.

[18] M. C. N. Cheng, J. F. R. Duncan, S. M. Harrison, J. A. Harvey, S. Kachru, and B. C. Rayhaun, "Attractive Strings and Five-Branes, Skew-Holomorphic Jacobi Forms and Moonshine," 2017.

[19] J. M. Maldacena, A. Strominger, and E. Witten, "Black hole entropy in M theory," *JHEP*, vol. 12, p. 002, 1997.

[20] D. Zagier, "Ramanujan's mock theta functions and their applications (after Zwegers and Ono-Bringmann)," *Astérisque*, no. 326, pp. Exp. No. 986, vii–viii, 143–164 (2010), 2009. Séminaire Bourbaki. Vol. 2007/2008.

[21] S. Zwegers, "Mock theta functions," July 2008.

[22] M. C. N. Cheng and J. F. R. Duncan, "Mathieu Moonshine at Dual Weight," *In preparation. See also my talk in the Conference String–Math 2017, 28 July 2017*.

[23] C. Vafa and E. Witten, "A strong coupling test of S-duality," *Nuclear Phys. B*, vol. 431, no. 1-2, pp. 3–77, 1994.

[24] J. Troost, "The non-compact elliptic genus: mock or modular," *JHEP*, vol. 1006, p. 104, 2010.

[25] S. Alexandrov, J. Manschot, and B. Pioline, "D3-instantons, Mock Theta Series and Twistors," *JHEP*, vol. 04, p. 002, 2013.

[26] M. C. N. Cheng and S. Harrison, "Umbral Moonshine and K3 Surfaces," *Commun. Math. Phys.*, vol. 339, no. 1, pp. 221–261, 2015.

[27] T. Eguchi and Y. Sugawara, "Non-holomorphic Modular Forms and SL(2,R)/U(1) Superconformal Field Theory," *JHEP*, vol. 03, p. 107, 2011.

[28] T. Eguchi and Y. Sugawara, "SL(2,R) / U(1) supercoset and elliptic genera of noncompact Calabi-Yau manifolds," *JHEP*, vol. 05, p. 014, 2004.

[29] S. K. Ashok, N. Doroud, and J. Troost, "Localization and real Jacobi forms," *JHEP*, vol. 04, p. 119, 2014.

[30] S. K. Ashok and J. Troost, "A Twisted Non-compact Elliptic Genus," *JHEP*, vol. 03, p. 067, 2011.

[31] S. Murthy, "A holomorphic anomaly in the elliptic genus," *JHEP*, vol. 06, p. 165, 2014.

[32] M. C. N. Cheng and E. Verlinde, "Dying Dyons Don't Count," *JHEP*, vol. 09, p. 070, 2007.

[33] V. G. Kac and M. Wakimoto, "Representations of affine superalgebras and mock theta functions II," *Adv. Math.*, vol. 300, pp. 17–70, 2016.

[34] S. Ramanujan, *Collected papers of Srinivasa Ramanujan.* AMS Chelsea Publishing, Providence, RI, 2000. Edited by G. H. Hardy, P. V. Seshu Aiyar and B. M. Wilson, Third printing of the 1927 original, With a new preface and commentary by Bruce C. Berndt.

[35] D. Zagier, “Nombres de classes et formes modulaires de poids 3/2,” *C. R. Acad. Sci. Paris Sér. A-B*, vol. 281, no. 21, pp. Ai, A883–A886, 1975.

[36] K. Yoshioka, “The Betti numbers of the moduli space of stable sheaves of rank 2 on $\mathbf{P}^2$,” *J. Reine Angew. Math.*, vol. 453, pp. 193–220, 1994.

[37] C. Vafa and E. Witten, “A Strong coupling test of S duality,” *Nucl. Phys.*, vol. B431, pp. 3–77, 1994.

[38] P. DiFrancesco, *Conformal field theory / Philippe Di Francesco, Pierre Mathieu, David Snchal.* Graduate texts in contemporary physics, New York: Springer, 1997.

[39] R. Blumenhagen, *Introduction to Conformal Field Theory With Applications to String Theory / / by Ralph Blumenhagen, Erik Plauschinn.* Lecture Notes in Physics, 779, Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2009.

[40] T. Gannon, “Monstrous moonshine: the first twenty-five years,” *Bull. London Math. Soc.*, vol. 38, no. 1, pp. 1–33, 2006.

[41] M. R. Gaberdiel, “An introduction to conformal field theory,” *Reports on Progress in Physics*, vol. 63, pp. 607–667, April 2000.

[42] X. Yin, “Tasi lectures on 2d cfts,” 2017. Lecture notes for TASI 2017.

[43] G. Felder, “The kzb equations on riemann surfaces,” September 1996.

[44] J. Fuchs, *Affine Lie algebras and quantum groups : an introduction, with applications in conformal field theory / Jrgen Fuchs.* Cambridge monographs on mathematical physics, Cambridge [etc.: Cambridge University Press, 1992.

[45] T. Kohno, *Conformal field theory and topology / Toshitake Kohno ; [translated from the Japanese by the author].* Translations of mathematical monographs v. 210, Providence, RI: American Mathematical Society, 2002.

[46] Y. Zhu, “Modular invariance of characters of vertex operator algebras,” *Journal of the American Mathematical Society*, vol. 9, no. 1, pp. 237–302, 1996.

[47] A. Cappelli, C. Itzykson, and J. Zuber, “The ADE Classification of Minimal and A1(1) Conformal Invariant Theories,” *Commun.Math.Phys.*, vol. 113, p. 1, 1987.

[48] T. Gannon, “The classification of SU(3) modular invariants revisited,” *Ann. Inst. H. Poincaré Phys. Théor.*, vol. 65, no. 1, pp. 15–55, 1996.

[49] R. Dijkgraaf, C. Vafa, E. P. Verlinde, and H. L. Verlinde, “The Operator Algebra of Orbifold Models,” *Commun. Math. Phys.*, vol. 123, p. 485, 1989.

[50] T.-L. Costache, “On irreducible projective representations of finite groups,” *Surveys in Mathematics and its Applications*, vol. 4, pp. 191–214, December 2009.

[51] C. Dong, H. Li, and G. Mason, “Modular invariance of trace functions in orbifold theory and generalized Moonshine,” *Communications in Mathematical Physics*, vol. 214, pp. 1–56, 2000.

[52] P. Bantay, “Orbifolds and hopf algebras,” *Physics Letters B*, vol. 245, no. 3, pp. 477–479, 1990.

[53] R. Dijkgraaf, V. Pasquier, and P. Roche, “QuasiHopf algebras, group cohomology and orbifold models,” *Nucl. Phys. Proc. Suppl.*, vol. 18B, pp. 60–72, 1990.

[54] S. Ochanine *Topology. An International Journal of Mathematics*, vol. 26, p. 143, 1987.

[55] E. Witten, “ELLIPTIC GENERA AND QUANTUM FIELD THEORY,” *Commun. Math. Phys.*, vol. 109, p. 525, 1987.

[56] P. S. Landweber, ed., *Elliptic curves and modular forms in algebraic topology.* Springer-Verlag, Berlin, 1988.

[57] E. Witten, “On the landau-ginzburg description of $n = 2$ minimal models,” *Int.J.Mod.Phys.A*, vol. 9:4783-4800,1994, Int.J.Mod.Phys.A9:4783-4800,1994.

[58] T. Kawai, Y. Yamada, and S.-K. Yang, “Elliptic genera and n=2 superconformal field theory,” *Nucl.Phys.B*, vol. 414:191-212,1994, Nucl.Phys.B414:191-212,1994.

[59] A. Hanany, N. Prezas, and J. Troost, “The Partition function of the two-dimensional black hole conformal field theory,” *JHEP*, vol. 04, p. 014, 2002.

[60] J. H. Bruinier and J. Funke, “On two geometric theta lifts,” *Duke Math. J.*, vol. 125, no. 1, pp. 45–90, 2004.

[61] G. Mason, “Vertex operator algebras, modular forms and moonshine,” vol. 8, pp. 335–365, 2014.

[62] J. F. R. Duncan, M. J. Griffin, and K. Ono, “Moonshine,” *Res. Math. Sci.*, vol. 2, pp. Art. 11, 57, 2015.

[63] J. H. Conway and S. P. Norton, “Monstrous Moonshine,” *Bull. London Math. Soc.*, vol. 11, p. 308 339, 1979.

[64] J. G. Thompson, “Some numerology between the Fischer-Griess Monster and the elliptic modular function,” *Bull. London Math. Soc.*, vol. 11, no. 3, pp. 352–353, 1979.

[65] J. G. Thompson, “Finite groups and modular functions,” *Bull. London Math. Soc.*, vol. 11, no. 3, pp. 347–351, 1979.

[66] A. P. Ogg, “Automorphismes de courbes modulaires,” p. 8, 1975.

[67] J. F. R. Duncan and I. B. Frenkel, “Rademacher sums, moonshine and gravity,” *Commun. Number Theory Phys.*, vol. 5, no. 4, pp. 1–128, 2011.

[68] P. Fong, “Characters arising in the monster-modular connection,” in *Proceedings of Symposia in Pure Mathematics*, vol. 37, pp. 557–559, AMER MATHEMATICAL SOC 201 CHARLES ST, PROVIDENCE, RI 02940-2213, 1980.

[69] S. D. Smith, “On the head characters of the monster simple group,” *Contemp. Math.*, vol. 45, pp. 303–313, 1985.

[70] I. B. Frenkel, J. Lepowsky, and A. Meurman, "A natural representation of the Fischer-Griess Monster with the modular function J as character," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 81, no. 10, Phys. Sci., pp. 3256–3260, 1984.

[71] I. B. Frenkel, J. Lepowsky, and A. Meurman, "A moonshine module for the Monster," in *Vertex operators in mathematics and physics (Berkeley, Calif., 1983)*, vol. 3 of *Math. Sci. Res. Inst. Publ.*, pp. 231–273, New York: Springer, 1985.

[72] R. L. Griess, Jr., "The friendly giant," *Invent. Math.*, vol. 69, no. 1, pp. 1–102, 1982.

[73] S. Norton, "Generalized Moonshine," *Proc. Symp. Pure Math*, vol. 47, pp. 208–209, 1987.

[74] L. Dixon, P. Ginsparg, and J. Harvey, "Beauty and the beast: Superconformal symmetry in a monster module," *Communications in Mathematical Physics*, vol. 119, pp. 221–241, June 1988.

[75] S. Carnahan, "Generalized moonshine iv: Monstrous lie algebras," August 2012.

[76] J. F. Duncan, "Super-moonshine for conway's largest sporadic group," *Duke Math. J.*, vol. 139, pp. 255–315, August 2007.

[77] J. f. r. Duncan and S. Mack-Crane, "The moonshine module for conways group," *Forum of Mathematics, Sigma*, vol. 3, December 2015.

[78] J. F. R. Duncan and S. Mack-Crane, "Derived equivalences of K3 surfaces and twined elliptic genera," *Res. Math. Sci.*, vol. 3, pp. Paper No. 1, 47, 2016.

[79] M. C. N. Cheng, X. Dong, J. F. R. Duncan, S. Harrison, S. Kachru, and T. Wrase, "Mock Modular Mathieu Moonshine Modules," 2014. [Res. Math. Sci.2,13(2015)].

[80] S. M. Harrison, N. M. Paquette, and R. Volpato, "A Borcherds-Kac-Moody superalgebra with Conway symmetry," 2018.

[81] T. Eguchi, H. Ooguri, and Y. Tachikawa, "Notes on the K3 Surface and the Mathieu group M_{24}," *Exper.Math.*, vol. 20, pp. 91–96, 2011.

[82] T. Eguchi and K. Hikami, "Superconformal Algebras and Mock Theta Functions 2. Rademacher Expansion for K3 Surface," *Communications in Number Theory and Physics*, vol. 3,, pp. 531–554, Apr. 2009.

[83] M. C. N. Cheng, "$K3$ Surfaces, $N = 4$ Dyons, and the Mathieu Group M_{24}," May 2010.

[84] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu twining characters for K3," *JHEP*, vol. 1009, p. 058, 2010.

[85] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu Moonshine in the elliptic genus of K3," *JHEP*, vol. 1010, p. 062, 2010.

[86] T. Eguchi and K. Hikami, "Note on Twisted Elliptic Genus of K3 Surface," *Phys.Lett.*, vol. B694, pp. 446–455, 2011.

[87] M. C. N. Cheng and J. F. R. Duncan, “The largest Mathieu group and (mock) automorphic forms,” vol. 85, pp. 53–82, 2012.

[88] T. Gannon, “Much ado about Mathieu,” 2012.

[89] M. C. Cheng, J. F. Duncan, and J. A. Harvey, “Umbral Moonshine,” *Commun.Num.TheorPhys.*, vol. 08, pp. 101–242, 2014.

[90] M. C. N. Cheng, J. F. R. Duncan, and J. A. Harvey, “Umbral Moonshine and the Niemeier Lattices,” 2013.

[91] M. C. N. Cheng and J. F. R. Duncan, “On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine,” Oct. 2011.

[92] M. C. N. Cheng, J. F. R. Duncan, and J. A. Harvey, “Weight One Jacobi Forms and Umbral Moonshine,” 2017.

[93] J. F. Duncan, M. J. Griffin, and K. Ono, “Proof of the Umbral Moonshine Conjecture,”

[94] J. F. R. Duncan and J. A. Harvey, “The Umbral Moonshine Module for the Unique Unimodular Niemeier Root System,” 2014.

[95] J. F. R. Duncan and A. O’Desky, “Super Vertex Algebras, Meromorphic Jacobi Forms and Umbral Moonshine,” 2017.

[96] M. C. N. Cheng and J. F. R. Duncan, “Meromorphic Jacobi Forms of Half-Integral Index and Umbral Moonshine Modules,” 2017.

[97] V. Anagiannis, M. C. N. Cheng, and S. M. Harrison, “K3 Elliptic Genus and an Umbral Moonshine Module,” 2017.

[98] M. C. N. Cheng, P. de Lange, and D. P. Z. Whalen, “Generalised Umbral Moonshine,” 2016.

[99] J. A. Harvey and B. C. Rayhaun, “Traces of Singular Moduli and Moonshine for the Thompson Group,” *Commun. Num. Theor. Phys.*, vol. 10, pp. 23–62, 2016.

[100] J. F. Duncan, J. A. Harvey, and R. B. C., “Skew-Holomorphic Jacobi Forms and Moonshine,” *In preparation.*

[101] D. Zagier, “Traces of singular moduli,” in *Motives, polylogarithms and Hodge theory, Part I (Irvine, CA, 1998)*, vol. 3 of *Int. Press Lect. Ser.*, pp. 211–244, Int. Press, Somerville, MA, 2002.

[102] R. Borcherds, “Automorphic forms on o s +2,2(r) and infinite products,” *Inventiones mathematicae*, vol. 120, pp. 161–213, December 1995.

[103] S. Carnahan, “Generalized moonshine, II: Borcherds products,” *Duke Math. J.*, vol. 161, no. 5, pp. 893–950, 2012.

[104] M. Griffin and M. Mertens, “A proof of the thompson moonshine conjecture,” *Research in the Mathematical Sciences*, vol. 3, pp. 1–32, December 2016.

[105] J. F. Duncan, M. H. Mertens, and K. Ono, "O'Nan Moonshine and Arithmetic,"

[106] M. C. N. Cheng, J. F. Duncan, Mertens, and M. H., "Class Number Moonshine," *In preparation. See also my talk in the Conference on Number Theory, Geometry, Moonshine & Strings, 08 Sep 2017.*

[107] W. Kohnen, "Fourier coefficients of modular forms of half-integral weight," *Math. Ann.*, vol. 271, no. 2, pp. 237–268, 1985.

[108] B. H. Gross and D. B. Zagier, "Heegner points and derivatives of L-series," *Invent. Math.*, vol. 84, no. 2, pp. 225–320, 1986.

[109] V. A. Kolyvagin, "Finiteness of $E(\mathbf{Q})$ and $\mathrm{SH}(E, \mathbf{Q})$ for a subclass of Weil curves," *Izv. Akad. Nauk SSSR Ser. Mat.*, vol. 52, no. 3, pp. 522–540, 670–671, 1988.

[110] C. Skinner and E. Urban, "The Iwasawa main conjectures for GL_2," *Invent. Math.*, vol. 195, no. 1, pp. 1–277, 2014.

[111] N. M. Paquette, D. Persson, and R. Volpato, "Monstrous bps-algebras and the superstring origin of moonshine," January 2016.

[112] N. M. Paquette, D. Persson, and R. Volpato, "BPS Algebras, Genus Zero, and the Heterotic Monster," *J. Phys.*, vol. A50, no. 41, p. 414001, 2017.

[113] S. Carnahan, "Generalized moonshine i: Genus zero functions," *Algebra and Number Theory*, vol. 4:6, pp. 649–679, Dec. 2008.

[114] D. Persson and R. Volpato, "Fricke S-duality in CHL models," *JHEP*, vol. 12, p. 156, 2015.

[115] A. Strominger and C. Vafa, "Microscopic origin of the bekenstein-hawking entropy," *Phys.Lett.B*, vol. 379:99-104,1996, Phys.Lett.B379:99-104,1996.

[116] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," *Adv.Theor.Math.Phys.*, vol. 2, pp. 231–252, 1998.

[117] T. Eguchi and A. Taormina, "Unitary representations of the $N = 4$ superconformal algebra," *Phys. Lett. B*, vol. 196, no. 1, pp. 75–81, 1987.

[118] T. Eguchi and A. Taormina, "Character formulas for the $N = 4$ superconformal algebra," *Phys. Lett. B*, vol. 200, no. 3, pp. 315–322, 1988.

[119] T. Eguchi and A. Taormina, "On the unitary representations of $N = 2$ and $N = 4$ superconformal algebras," *Phys. Lett. B*, vol. 210, no. 1-2, pp. 125–132, 1988.

[120] M. C. N. Cheng and S. Harrison, "Umbral Moonshine and K3 Surfaces," *Commun. Math. Phys.*, vol. 339, no. 1, pp. 221–261, 2015.

[121] H. Ooguri and C. Vafa, "Two-dimensional black hole and singularities of CY manifolds," *Nucl. Phys.*, vol. B463, pp. 55–72, 1996.

[122] J. A. Harvey, S. Lee, and S. Murthy, "Elliptic genera of ALE and ALF manifolds from gauged linear sigma models," *JHEP*, vol. 02, p. 110, 2015.

[123] A. Taormina and K. Wendland, "The Conway Moonshine Module is a Reflected K3 Theory," 2017.

[124] T. Creutzig, J. F. R. Duncan, and W. Riedler, "Self-Dual Vertex Operator Superalgebras and Superconformal Field Theory," *J. Phys.*, vol. A51, no. 3, p. 034001, 2018.

[125] M. C. N. Cheng, S. M. Harrison, R. Volpato, and M. Zimet, "K3 string theory, lattices and moonshine," December 2016.

[126] S. Mukai, "Finite groups of automorphisms of $K3$ surfaces and the Mathieu group," *Invent. Math.*, vol. 94, no. 1, pp. 183–221, 1988.

[127] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Symmetries of $K3$ sigma models," June 2011.

[128] S. Kondō, "Niemeier lattices, Mathieu groups, and finite groups of symplectic automorphisms of $K3$ surfaces," *Duke Math. J.*, vol. 92, no. 3, pp. 593–603, 1998. With an appendix by Shigeru Mukai.

[129] D. Huybrechts, "On derived categories of K3 surfaces, symplectic automorphisms and the Conway group," vol. 69, pp. 387–405, 2016.

[130] M. C. N. Cheng, J. F. R. Duncan, S. M. Harrison, and S. Kachru, "Equivariant K3 Invariants," *Commun. Num. Theor. Phys.*, vol. 11, pp. 41–72, 2017.

[131] P. S. Aspinwall, "Enhanced gauge symmetries and K3 surfaces," *Phys. Lett.*, vol. B357, pp. 329–334, 1995.

[132] E. Witten, "String theory dynamics in various dimensions," *Nucl. Phys.*, vol. B443, pp. 85–126, 1995.

[133] P. S. Aspinwall, "K3 surfaces and string duality," in *Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996*, pp. 421–540, 1996.

[134] W. Nahm and K. Wendland, "A Hiker's guide to K3: Aspects of N=(4,4) superconformal field theory with central charge c = 6," *Commun. Math. Phys.*, vol. 216, pp. 85–138, 2001.

[135] N. M. Paquette, R. Volpato, and M. Zimet, "No More Walls! A Tale of Modularity, Symmetry, and Wall Crossing for 1/4 BPS Dyons," *JHEP*, vol. 05, p. 047, 2017.

[136] A. Taormina and K. Wendland, "The overarching finite symmetry group of Kummer surfaces in the Mathieu group M_{24}," *JHEP*, vol. 08, p. 125, 2013.

[137] A. Taormina and K. Wendland, "A twist in the M24 moonshine story," 2013.

[138] M. C. N. Cheng, F. Ferrari, S. M. Harrison, and N. M. Paquette, "Landau-Ginzburg Orbifolds and Symmetries of K3 CFTs," *JHEP*, vol. 01, p. 046, 2017.

[139] M. R. Gaberdiel, C. A. Keller, and H. Paul, "Mathieu moonshine and symmetry surfing," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, November 2017.

[140] M. C. N. Cheng, F. Ferrari, S. M. Harrison, and N. M. Paquette, "Landau-Ginzburg Orbifolds and Symmetries of K3 CFTs," 2015.

[141] D. Persson and R. Volpato, "Second Quantized Mathieu Moonshine," *Commun. Num. Theor. Phys.*, vol. 08, pp. 403–509, 2014.

[142] S. Kachru, N. M. Paquette, and R. Volpato, "3D String Theory and Umbral Moonshine," 2016.

[143] M. Zimet, "Umbral Moonshine and String Duality," 2018.

[144] M. C. Cheng, X. Dong, J. Duncan, J. Harvey, S. Kachru, *et al.*, "Mathieu Moonshine and N=2 String Compactifications," 2013.

[145] S. Harrison, S. Kachru, and N. M. Paquette, "Twining Genera of (0,4) Supersymmetric Sigma Models on K3," *JHEP*, vol. 04, p. 048, 2014.

[146] J. A. Harvey and S. Murthy, "Moonshine in Fivebrane Spacetimes," *JHEP*, vol. 01, p. 146, 2014.

[147] J. A. Harvey, S. Murthy, and C. Nazaroglu, "ADE Double Scaled Little String Theories, Mock Modular Forms and Umbral Moonshine," *JHEP*, vol. 05, p. 126, 2015.

What is the monster.

Richard E. Borcherds, *

Mathematics department, Evans Hall #3840, University of California at Berkeley, CA 94720-3840 U. S. A.

e-mail: reb@math.berkeley.edu

www home page www.math.berkeley.edu/˜reb

When I was a graduate student, my supervisor John Conway would bring his 1 year old son into the department, who was soon known as the baby monster. A more serious answer to the question is that the monster is the largest of the (known**) sporadic simple groups. Its name comes from its size: the number of elements is

$$\begin{aligned}&8080, 17424, 79451, 28758, 86459, 90496, 17107, 57005, 75436, 80000, 00000\\=&2^{46}.3^{20}.5^{9}.7^{6}.11^{2}.13^{3}.17.19.23.29.31.41.47.59.71,\end{aligned}$$

about equal to the number of elementary particles in the planet Jupiter.

The monster was originally predicted to exist by Fischer and by Griess in the early 1970's. Griess constructed it a few years later in an extraordinary tour de force as the group of linear transformations on a vector space of dimension 196883 that preserved a certain commutative but non-associative bilinear product, now called the Griess product.

Our knowledge of the structure and representations of the monster is now pretty good. The 194 irreducible complex representations were worked out by Fischer, Livingstone and Thorne (before the monster was even shown to exist) They take up 8 large pages in the atlas [A] of finite groups, which is the best single source of information about the monster (and other finite simple groups). The subgroup structure is mostly known; in particular there is an almost complete list of the maximal subgroups, and the main gaps in our knowledge concern embeddings of very small simple groups in the monster. If anyone wishes to multiply elements of the monster explicitly, R. A. Wilson can supply two matrices that generate the monster. But there is a catch: each matrix takes up about 5 gigabytes of storage, and to quote from Wilson's atlas page: "standard generators have now been made as 196882×196882 matrices over GF(2)... They have been multiplied together, using most of the computing resources of Lehrstuhl D für Mathematik, RWTH Aachen for about 45 hours...". (The difficulty of multiplying two elements of the monster is caused not so much by its huge size as by the lack of "small" representations; for example, the symmetric group S_{50} is quite a lot bigger than the monster, but it only takes a few minutes to multiply two elements by hand.) Finally the modular representations of the monster for large primes were worked out by Hiss and Lux; the ones for small primes still seem to be out of reach at the moment.

---

* Supported by NSF grant DMS-9970611.

** The announcement of the classification of the finite simple groups about 20 years ago was a little over-enthusiastic, but a recent 1300 page preprint by Aschbacher and Smith should finally complete it.

In the late 1970's John McKay decided to switch from finite group theory to Galois theory. One function that turns up in Galois theory is the elliptic modular function

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + \cdots = \sum c(n)q^n$$

($q = e^{2\pi i\tau}$), which is essentially the simplest non-constant function invariant under the action $\tau \mapsto (a\tau + b)/(c\tau + d)$ of $SL_2(\mathbf{Z})$ on the upper half plane $\{\tau | \Im(\tau) > 0\}$. He noticed that the coefficient 196884 of q^1 was almost equal to the degree 196883 of the smallest complex representation of the monster (up to a small experimental error). The term "moonshine" roughly means weird relations between sporadic groups and modular functions (and anything else) similar to this. It was clear to many people that this was just a meaningless coincidence; after all, if you have enough large integers from various areas of mathematics then a few are going to be close just by chance, and John McKay was told that his observation was about as useful as looking at tea-leaves. John Thompson took McKay's observation further, and pointed out that the next few coefficients of the elliptic modular function were also simple linear combinations of dimensions of irreducible representations of the monster; for example, $21493760 = 21296876 + 196883 + 1$. He suggested that there should be a natural infinite dimensional graded representation $V = \sum_{n\in\mathbf{Z}} V_n$ of the monster such that the dimension of V_n is the coefficient $c(n)$ of q^n in $j(\tau)$, at least for $n \neq 0$. (The constant term of $j(\tau)$ is arbitrary as adding a constant to j still produces a function invariant under $SL_2(\mathbf{Z})$, and is set equal to 744 mainly for historical reasons.) Conway and Norton [C-N] followed up Thompson's suggestion of looking at the McKay-Thompson series $T_g(\tau) = \sum_n Trace(g|V_n)q^n$ whose coefficients are given by the traces of elements g of the monster on the representations V_n, and found by calculating the first few terms that they all seemed to be Hauptmoduls of genus 0. (A Hauptmodul is a function similar to j, but invariant under some group other than $SL_2(\mathbf{Z})$). Atkin, Fong, and Smith showed by computer calculation that there was indeed an infinite dimensional graded representation of the monster whose McKay-Thompson series were the Hauptmoduls found by Conway and Norton, and soon afterwards Frenkel, Lepowsky, and Meurman explicitly constructed this representation using vertex operators.

If a group acts on a vector space it is natural to ask if it preserves any algebraic structure, such as a bilinear form or product. The monster module constructed by FLM has a vertex algebra structure invariant under the action of the monster. Unfortunately there is no easy way to explain what a vertex algebra is; see [K] for the best introduction to them. Vertex algebras are a generalization of commutative rings with derivations (at least in characteristic 0). Roughly speaking they can be thought of as commutative rings with derivation where the ring multiplication is not quite defined everywhere; this is analogous to rational maps in algebraic geometry, which are also not quite defined everywhere. A more concrete but less intuitive definition of a vertex algebra is that it consists of a space with a countable number of bilinear products satisfying certain rather complicated identities. In the case of the monster vertex algebra $V = \oplus V_n$, this gives bilinear maps from $V_i \times V_j$ to V_k for all integers i, j, k, and the special case of the map from $V_2 \times V_2$ to V_2 is (essentially) the Griess product. So the Griess algebra is a sort of section of the monster vertex algebra.

Following an idea of Frenkel, the monster vertex algebra can be used to construct the monster Lie algebra by using the Goddard-Thorn no-ghost theorem from string theory.

This is a $\mathbf{Z}^2$-graded Lie algebra, whose piece of degree $(m,n) \in \mathbf{Z}^2$ has dimension $c(mn)$ whenever $(m,n) \neq (0,0)$. The monster should be thought of as a group of "diagram automorphisms" of this Lie algebra, in the same way that the symmetric group S_3 is a group of diagram automorphisms of the Lie algebra D_4. The monster Lie algebra has a denominator formula, similar to the Weyl denominator formula for finite dimensional Lie algebras and the Macdonald-Kac identities for affine Lie algebras, which looks like

$$j(\sigma) - j(\tau) = p^{-1} \prod_{\substack{m>0 \\ n\in\mathbf{Z}}} (1 - p^m q^n)^{c(mn)}$$

where $p = e^{2\pi i\sigma}$, $q = e^{2\pi i\tau}$. This formula was discovered independently in the 1980's by several people, including Koike, Norton, and Zagier. There are similar identities with $j(\tau)$ replaced by the McKay-Thompson series of any element of the monster, and Cummins and Gannon showed that any function satisfying such identities is a Hauptmodul. So this provides some sort of explanation of Conway and Norton's observation that the McKay-Thompson series are all Hauptmoduls.

So the question "What is the monster" now has several reasonable answers:

1 The monster is the largest sporadic simple group, or alternatively the unique simple group with its order.
2 It is the automorphism group of the Griess algebra.
3 It is the automorphism group of the monster vertex algebra. (This is probably the best answer.)
4 It is a group of diagram automorphisms of the monster Lie algebra.

Unfortunately none of these definitions is completely satisfactory. At the moment all constructions of the algebraic structures above seem artificial; they are constructed as sums of two or more apparently unrelated spaces, and it takes a lot of effort to define the algebraic structure on the sum of these spaces and to check that the monster acts on the resulting structure.

References.

[A] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, R. A. Wilson, Atlas of finite groups, Clarendon Press, Oxford, 1985.

[C-N] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[K] V. Kac, Vertex algebras for beginners. Second edition. University Lecture Series, 10. American Mathematical Society, Providence, RI, 1998.

The j-Function and the Monster

Asa Scherer

August 16, 2010

Chapter 1

Introduction

In a set of Hauptmoduls lived a function by the name of Jay. The year was 1979. Jay was quite old, though whether he was 100 or had lived forever was a matter only mathematical philosophers cared about. This particular day, Jay went about his usual business of generating other modular functions. Jay had always been connected to many of his mathematical-object friends. Little did anyone know, Jay held a dark secret. He was also connected... to a Monster.

Now that I have the attention of all my readers aged 5 to 10 and no one else, let's delve into some advanced mathematics!

In this ridiculous introduction, our fellow Jay is standing in for the Klein j-invariant, a complex-valued function that has many fascinating properties, not least of all that any modular function can be written as a rational function of the j-function. The mysterious Monster is the Fischer-Greiss Monster group, the largest of the sporadic simple groups. The real protagonists at the heart of this paper are John Conway and Simon Norton, who formulated the conjecture connecting these two seemingly disparate objects, and Richard Borcherds, who proved [Bor92] their conjecture in 1992.

The first published notion that the j-function was in any way related to the Monster came in 1979, when Conway and Norton noted in [CN79] that each coefficient in the q-expansion of the j-function could be written as a (nontrivial) integral linear combination of the dimensions of irreducible representations of the Monster. For example, the first relevant coefficient, 196884, can be written $196883 + 1$, where 196883 is the dimension of the smallest non-trivial representation of the Monster, and 1 is the dimension of the trivial representation. This observation led to their conjecture, which they titled "monstrous moonshine." This is a rather mysterious and evocative name for a mathematical conjecture, meant to insinuate that it is "distilling information illegally" [Gan04] from the Monster. It also evokes a sense of absurdity, that this connection is so far-fetched that it must be just a coincidence. This, of course, is not the case. Monstrous moonshine indeed illustrates a profound relationship between the j-function. In fact, the conjecture goes beyond just the j-function, but the j-function serves as an especially illustrative case of the conjecture.

Before we proceed any further, let us meet our main characters.

1.1 The Monster

Don't let the Monster group's name fool you! It's really quite a friendly group. Unless you're trying to analyze any piece of it, really, in which case it is quite a monster.

Before we define what the monster really is, we recall some group theoretical definitions:

Definition 1.1.1. *A* **simple group** G *is a group with no nontrivial normal subgroups.*

Normal subgroups, of course, are subgroups H where $ghg^{-1} \in H$ for all $g \in G$ and for all $h \in H$. By nontrivial we mean subgroups that are not the whole group or the trivial subgroup.

Simple groups are important to the study of groups, as any finite group can be decomposed into a composition series:

Definition 1.1.2. *A* **composition series** *is an inclusion chain of normal subgroups*

$$1 = H_0 \triangleleft H_1 \triangleleft H_2 \triangleleft H_3 \triangleleft \cdots \triangleleft G,$$

where H_{i+1}/H_i *is simple for all* $i \geq 1$.

The Jordan-Holder theorem states that any such series is unique up to permutation and isomorphism. So, in essence, simple groups can be considered the "building blocks" of all finite groups, analogous to prime numbers in number theory.

The classification of all finite simple groups was a massive undertaking, only completed in 1983 [Sol01]. The results of this achievement give us the following complete list of finite simple groups:

Infinite families:

1. The cyclic groups $\mathbb{Z}/p\mathbb{Z}$, p prime
2. The alternating groups A_n, $n \geq 5$
3. 16 infinite families of Lie type.

The rest of the simple groups are:

4. Exactly 26 so-called "sporadic" groups.

The **Monster group** $\mathbb{M}$ is the largest of these sporadic simple groups, with group order

$$|\mathbb{M}| = 2^{45} \cdot 3^{20} \cdot 5^{9} \cdot 7^{6} \cdot 11^{2} \cdot 13^{3} \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \text{[Gan04]}.$$

So... It's huge. And yet not one of its many nontrivial subgroups is normal.

The Monster was postulated to exist in 1973 by Bernd Fischer and Robert Griess, though it was not explicitly constructed until 1980, when Griess constructed it as the group of automorphisms of a particular 196883-dimensional algebra. The Monster has since been built using other methods. [Gan04] The Monster is a very complicated group, having 194 conjugacy classes, and it contains 19 of the other sporadic simple groups as subgroups or quotients of subgroups, which include the adorably-named Baby Monster. (See [Gan04] and [Sol01] for more information.)

1.2 The j-function

Definition 1.2.1. *The j-function is a complex-valued function defined on all $\tau \in \mathbb{C}$ such that*

$$j(\tau) = 1728 \cdot \frac{g_2(\tau)^3}{g_2(\tau)^3 - 27g_3(\tau)^2},$$

where $g_2(\tau)$ and $g_3(\tau)$ are certain infinite sums over all the points in a particular lattice, as defined in Section 2.4.

Jay has many guises, though! Setting $q = e^{2\pi i\tau}$, we can write

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 + \cdots := \frac{1}{q} + \sum_{n=0}^{\infty} c(n)q^n,$$

as seen in [Gan04]. This expression is called the *q-expansion* of j, and it is in these integer coefficients that we find the fascinating connection with the Monster. For instance, Conway and Norton's original paper presenting monstrous moonshine [CN79] noted that

$$\begin{aligned} 196884 &= 196883 + 1 \\ 21493760 &= 21296876 + 196883 \\ 864299970 &= 842609326 + 2 \cdot 21296876 + 2 \cdot 196883 + 2 \cdot 1, \end{aligned}$$

where the numbers on the right are dimensions of irreducible representations of the Monster. (We define representations in Section 2.1.)

The j-function has quite a few rather remarkable properties on its own, some of which we will explore in Chapter 3. Namely, we will focus on proving two main theorems. First, though, some more definitions are in order:

Definition 1.2.2. *The group $SL_2(\mathbb{Z})$ is defined as the multiplicative group of all two-by-two matrices with integer entries and determinant 1; that is,*

$$SL_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1 \right\}.$$

Definition 1.2.3. *We write $SL_2(\mathbb{Z})\backslash\mathbb{H}$ to denote the quotient of the upper half-plane $\mathbb{H}$ by the action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$ (which will be defined in Chapter 2).*

We have two main goals in this paper. First, we will focus on proving the following properties of $j(\tau)$:

Theorem 1.2.4. *The j-function is a bijection between $SL_2(\mathbb{Z})\backslash\mathbb{H}$ and $\mathbb{C}$.*

Theorem 1.2.5. *Every modular function is a rational function of $j(\tau)$.*

Secondly, we will give the reader context and understanding of the actual statement of monstrous moonshine. We will first establish some context for these theorems by presenting some preliminary material in Chapter 2, including a brief introduction to representation theory, lattices, and modular forms. Chapter 3, following the logical flow laid out in [Cox89], covers the proof of our two main theorems, Theorems 1.2.4 and 1.2.5, and an additional property of $j(\tau)$ [Bor92] regarding the denominator formula. Chapter 4 defines the moonshine module, necessary to understand monstrous moonshine, and then states Borcherds' theorem, concluding by touching on a few of the related areas of moonshine research beyond the original conjecture.

A brief note: This paper assumes the reader to have a basic knowledge of group theory and complex analysis, but in general, we will be fairly explicit in defining our terms; many of these will be covered in Chapter Two.

Chapter 2

Preliminaries

Treat this chapter as a sort of smorgasbord of mathematical theory; a sampling of the delicious tools needed to understand the j-function and its Monstrous connection. We begin with an appetizer of Representation Theory, follow with a salad of Lattice, and finish with a Modular Function main course.

2.1 Representation Theory

Linear algebra is an incredibly useful tool, finding applications scattered throughout various mathematical and scientific disciplines. The beauty of linear algebra is that once we associate some sort of mathematical object to a matrix, we can instantly make many observations about the matrix, which then may apply to the original object. Group theorists in particular use linear algebra to simplify certain tasks of group structure analysis using maps called *representations*. In this section, we review some basic facts about representations of finite groups over $\mathbb{C}$. For more information, see [FH91].

Definition 2.1.1. *Given a finite group G, a* **representation of G on a vector space V over $\mathbb{C}$** *is a map ρ from G to the group of linear automorphisms of V, $GL(V)$, such that $\rho(g_1g_2) = \rho(g_1) \circ \rho(g_2)$.*

So, essentially, a representation transfers group properties into a linear algebraic context, where our linear transformations behave as their preimage group elements do. A note: We frequently refer to V itself as the representation, rather than the map ρ. Since our representations are vector spaces, we can certainly consider subspaces that are also representations.

Definition 2.1.2. *Let vector space V be a representation of group G. Let W be a subspace of V. If $\rho(g)(\mathbf{v}) \in W$ for all $\mathbf{v} \in W$, then we call W a* **subrepresentation** *of V with respect to G.*

As we have introduced the concept of subrepresentation, we wish to now define the most "basic" type of representation:

Definition 2.1.3. *An* **irreducible representation for** G *is a representation V that cannot be expressed as a direct sum of subrepresentations.*

In fact, we get a stronger result telling us that irreducible representations are our "representation building blocks":

Theorem 2.1.4. *For any representation V of G, we can express V as a unique direct sum of irreducible representations [FH91].*

Remark: Recall that if $A = (a_{ij})$ is a complex $n \times n$ matrix, then the trace of A is given by $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$, the sum of the diagonal entries in A.

Let $g \in G$. Then, if ρ is a representation for G on V, $\rho(g)$ is a linear transformation from V to itself. But, we know from linear algebra that any linear transformation can be expressed as a matrix, considering some standard, ordered arrangement of basis vectors. Hence we can talk about the trace of $\rho(g)$'s associated matrix. It is easily shown that this number is independent of choice of basis for V, which leads us to the following definition.

Definition 2.1.5. *Let $g \in G$, and let ρ be a representation for G on V. Let M_g be the matrix with respect to some ordered basis corresponding to $\rho(g)$. Then the trace of M_g is called the* **character of g with respect to the representation** *V, frequently denoted $\chi_V(g)$.*

Note that the character of the identity element $e \in G$ with respect to a representation V is always equal to the dimension of V, call it n, since $\rho(e)$ is the identity transformation, and thus its associated $n \times n$ matrix is just I_n, the $n \times n$ identity matrix. To get a better feel for what a representation really is, we present a brief example.

Example.

Let $G = S_3$ be the symmetric group of order 3, that is, the group of permutations of three elements. Let W be the vector space over $\mathbb{C}$ generated by the basis

$$B = \{(1,0,0), (0,1,0), (0,0,1)\}.$$

That is, $W = \mathbb{C}^3$. Define the representation map ρ like so: $\rho(g) = T_g$, where T_g is defined by

$$T_g(z_1, z_2, z_3) = (z_{g^{-1}(1)}, z_{g^{-1}(2)}, z_{g^{-1}(3)}),$$

where $g^{-1}(n)$ denotes the number (1, 2, or 3) that is sent to position n by the permutation g. Let's pick a specific g: Say, $g = (12)$, that is, the transposition that sends 1 to 2, 2 to 1, and 3 to itself. Then we can note where $\rho(g) = T_g$ sends our basis elements:

$$T_g(1,0,0) = (0,1,0)$$
$$T_g(0,1,0) = (1,0,0)$$
$$T_g(0,0,1) = (0,0,1)$$

Hence our matrix associated with T_g (with respect to our given ordered basis) is

$$M_g := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, the character of g with respect to representation W is

$$\mathrm{Tr}(M_g) = 0 + 0 + 1 = 1.$$

Note that any other permutation of order 2 in S_3 will also have character 1, and that any permutation of order 3 has a character of 0 (since none of the elements are fixed.) The identity element, of course, has character 3 as expected, since $dim(W) = 3$.

In fact, these observations can be correctly generalized by the following statement:

Theorem 2.1.6. *With respect to a given representation V of group G, any two elements of the same conjugacy class of G have the same character. [FH91]*

Remark: Thus, if we are interested in finding every group element's character, we only need to pick one representative from each conjugacy class. In fact, we have a further correlation between conjugacy classes and representations: The number of irreducible representations of a group G is equal to the number of congruence classes of G. [FH91]

Continuing our example, it turns out that our representation $W = \mathbb{C}^3$ is not actually irreducible, and can be expressed as the direct sum of a 1-dimensional and a 2-dimensional vector space. In particular, we can write $W = U \oplus V$, where $U = \mathrm{span}\{(1,1,1)\}$, and $V = \{(z_1, z_2, z_3) | z_1 + z_2 + z_3 = 0\}$. U is clearly irreducible, since it has dimension 1, and in fact V is irreducible, too [FH91]. Thus the direct sum $U \oplus V$ is the unique decomposition of $W = \mathbb{C}^3$ into irreducible representations of S_3.

2.2 Lattices

A **lattice** L is an additive subgroup of $\mathbb{C}$ generated by two $\mathbb{R}$-linearly independent complex numbers ω_1 and ω_2. That is, $L = \{m\omega_1 + n\omega_2 : m, n \in \mathbb{Z}\}$, with $\frac{\omega_1}{\omega_2} \notin \mathbb{R}$. This last requirement gives us the desired linear independence of ω_1 and ω_2. Geometrically, it is equivalent to saying that the two points do not lie on the same vector beginning at the origin, when viewing the complex plane as analogous to $\mathbb{R}^2$. Notationally, we write

$$L = [\omega_1, \omega_2] = \{m\omega_1 + n\omega_2 : m, n \in \mathbb{Z}\}.$$

Visually, lattices create nice parallelogram-tilings of $\mathbb{C}$. Two different lattices can be related in a particular way:

Definition 2.2.1. *The lattices $L = [\omega_1, \omega_2]$ and $L' = [\omega_1', \omega_2']$ are said to be* **homothetic** *if there exists $\lambda \in \mathbb{C}$ such that $L' = \lambda L$. That is, $[\omega_1', \omega_2'] = [\lambda\omega_1, \lambda\omega_2]$.*

Note that this property, homothety, defines an equivalence class on the set of all complex lattices; that is, we can look at the collection of all possible complex lattices and partition them by which lattices are homothetic to one another. We will call these equivalence classes **homothety classes**.

Now that we have lattices, we can define some important pieces of the j-function.

Definition 2.2.2. *Let $n \geq 3$, and let L be a lattice. Then*

$$G_n(L) = \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^n}$$

is called the **Eisenstein series of order n for L.**

You may now be asking, "What, if any, n do we know this expression to converge for?" Well, that's a good question, hypothetical reader, one which we will now answer:

Lemma 2.2.3. *Suppose L is a lattice and $s \in \mathbb{R}$ with $s > 2$. Then the series*

$$G_s(L) = \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^s}$$

converges absolutely.

Proof. Let ω_1, $\omega_2 \in \mathbb{C}$ be the generators of L, that is, $L = [\omega_1, \omega_2]$. So then we need to show that

$$\sum_{\omega \in L \setminus \{0\}} \frac{1}{|\omega|^s} = \sum_{m,n} \frac{1}{|m\omega_1 + n\omega_2|^s}$$

converges, where the expression on the right is a sum over all integer pairs $(m, n) \neq (0, 0)$.

Define

$$M = \min\{|x\omega_1 + y\omega_2| : x^2 + y^2 = 1\}.$$

Let x and y be real numbers such that at least one of x and y is nonzero. Note that

$$\left(\frac{x}{\sqrt{x^2+y^2}}\right)^2 + \left(\frac{y}{\sqrt{x^2+y^2}}\right)^2 = \frac{x^2+y^2}{x^2+y^2} = 1.$$

Hence, by the definition of M,

$$\left|\frac{x}{\sqrt{x^2+y^2}}\omega_1 + \frac{y}{\sqrt{x^2+y^2}}\omega_2\right| = \frac{|x\omega_1 + y\omega_1|}{\sqrt{x^2+y^2}} \geq M,$$

and so

$$|x\omega_1 + y\omega_2| \geq M\sqrt{x^2+y^2}$$

for all $x, y \in \mathbb{R}$.

We picked x and y so that at least one of them was nonzero to avoid having a zero denominator in our earlier expression, but this inequality is clearly true for $x = y = 0$, so it holds for all $x, y \in \mathbb{R}$.

But then

$$\sum_{m,n} \frac{1}{|m\omega_1 + n\omega_2|^s} \leq \frac{1}{M^s} \sum_{m,n} \frac{1}{(m^2+n^2)^{s/2}}.$$

Consider, then, the integral

$$\iint\limits_{x^2+y^2 \geq 1} \frac{1}{(x^2+y^2)^{s/2}} dx\ dy.$$

If this integral converges, we are done. Rewriting in terms of traditional polar coordinates r and θ, we have

$$\int\limits_{r=1}^{\infty}\int\limits_{\theta=0}^{2\pi}\frac{1}{r^s}r\,d\theta\,dr=\int\limits_{r=1}^{\infty}\int\limits_{\theta=0}^{2\pi}\frac{1}{r^{s-1}}d\theta\,dr.$$

Evaluating this expression with rudimentary calculus, we get

$$\int\limits_{r=1}^{\infty}\frac{2\pi}{r^{s-1}}dr.$$

Since $s>2$, then $s-1>1$, and so this integral exists and is a finite number. It follows that

$$\sum_{m,n}\frac{1}{|m\omega_1+n\omega_2|^s}$$

must converge.

$\square$

Hence, any Eisenstein series of order greater than 2 converges absolutely. Two particularly important Eisenstein series that we are especially interested in are

$$g_2(L)=60G_4(L)=60\sum_{\omega\in L\setminus\{0\}}\frac{1}{\omega^4}$$

and

$$g_3(L)=140G_6(L)=140\sum_{\omega\in L\setminus\{0\}}\frac{1}{\omega^6}.$$

We introduce now an important lattice-based function that will play a part in the proofs in Chapter 3.

Definition 2.2.4. *The* **Weierstrass $\wp$-function** *is defined by the following:*

$$\wp(z;L)=\frac{1}{z^2}+\sum_{\omega\in L\setminus\{0\}}\left(\frac{1}{(z-\omega)^2}-\frac{1}{\omega^2}\right)$$

for $z\in\mathbb{C}$ and L a lattice, where the sum on the right is a sum over all nonzero lattice points ω of L.

If it is clear (or doesn't matter) which lattice we are referring to, $\wp(z;L)$ can be abbreviated to simply $\wp(z)$. The following lemmas about the $\wp$-function can be found in [Cox89]:

Lemma 2.2.5. *The set of singularities of $\wp(z;L)$ is exactly equal to the set of lattice points in L.*

This lemma becomes fairly clear intuitively by looking at the denominator terms in $\wp(z;L)$.

Lemma 2.2.6. *The Laurent expansion of $\wp(z)$ for the lattice L takes the form*

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} p(g_2(L), g_3(L))z^{2n},$$

where $p(g_2(L), g_3(L))$ is a polynomial with rational coefficients, independent of L, in $g_2(L)$ and $g_3(L)$.

Proofs of these lemmas can be found in [Cox89]. Another function, which turns out to be an important piece of the j-function, is called the discriminant of the lattice L.

Definition 2.2.7. *The discriminant of the lattice L, denoted $\Delta(L)$, is given by*

$$\Delta(L) = g_2(L)^3 - 27g_3(L)^2.$$

Remark: As it just so happens, $\Delta(L)$ is the discriminant of the polynomial $p(x) = 4x^3 - g_2x - g_3$, which, incidentally, the Weirstrauss $\wp$-function satisfies like so, as mentioned by [Cox89]:

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3.$$

The roots of this polynomial $p(x)$ are $e_1 = \wp(\omega_1/2)$, $e_2 = \wp(\omega_2/2)$, and $e_3 = \wp(\frac{\omega_1+\omega_2}{2})$. In fact, these roots are distinct [Apo90]. Hence, since the discriminant of a polynomial is a product of differences of its roots, we get the following lemma:

Lemma 2.2.8. *For any lattice L, the discriminant $\Delta(L)$ is nonzero.*

2.3 Modular Functions

Modular functions: Is there anything more beautiful? Probably, but that won't stop this mathematician from waxing hyperbolic! But before we get to these awe-inspiring paragons of beauty, we must first introduce some context.

$\mathrm{SL}_2(\mathbb{Z})$, as noted before, is defined as the group, under multiplication, of 2×2 matrices with integer entries and determinant 1. In the context of modular functions, we call the group $\mathrm{SL}_2(\mathbb{Z})$ the **modular group** Γ. It turns out that $\mathrm{SL}_2(\mathbb{Z})$ acts on the upper half-plane $\mathbb{H} = \{a + bi \in \mathbb{C} : b > 0\}$ through linear fractional transformations like so:

If $\tau \in \mathbb{H}$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

An easy calculation shows that the action of $\mathrm{SL}_2(\mathbb{Z})$ takes elements of $\mathbb{H}$ to $\mathbb{H}$. We have a nice property of $\mathrm{SL}_2(\mathbb{Z})$ in that it is generated by two matrices: [Apo90]

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Note that

$$T \cdot \tau = \tau + 1 \quad \text{and} \quad S \cdot \tau = -\frac{1}{\tau}.$$

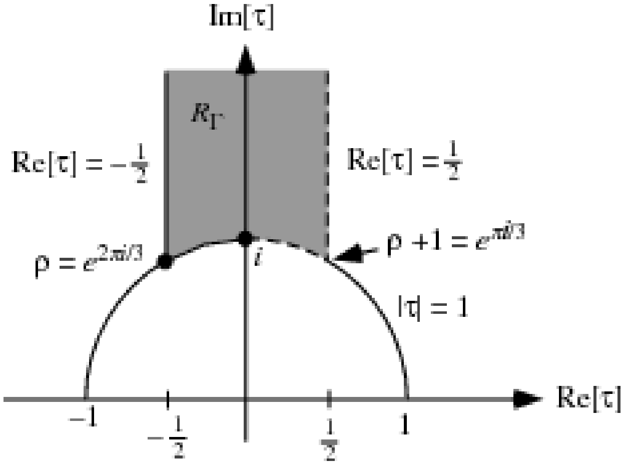


Figure 2.1: The Fundamental region F, here denoted R_Γ[BB87].

As we will see, modular forms have the property that they are invariant under $\mathrm{SL}_2(\mathbb{Z})$. Hence, we wish to analyze which points in $\mathbb{H}$ are "$\mathrm{SL}_2(\mathbb{Z})$-equivalent," that is, when two points $\tau, \tau' \in \mathbb{H}$ are such that $\tau' = \gamma \cdot \tau$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Specifically, is there a nice subregion of $\mathbb{H}$ where every point is $\mathrm{SL}_2(\mathbb{Z})$-distinct to every other point in the region and every $\mathrm{SL}_2(\mathbb{Z})$ equivalence class is represented? Why, yes! Yes, there is. This region is

$$F = \left\{ z \in \mathbb{C} : |z| > 1, -\frac{1}{2} \leq Re(z) < \frac{1}{2} \right\} \cup \left\{ z \in \mathbb{C} : |z| = 1, -\frac{1}{2} \leq Re(z) \leq 0 \right\},$$

and is called a fundamental region of the modular group Γ (see Figure 2.3).
F has the following properties:

1. Any two distinct points in F are not $\mathrm{SL}_2(\mathbb{Z})$-equivalent.
2. For any $\tau \in \mathbb{H}$, there is a point τ' in F such that τ and τ' are $\mathrm{SL}_2(\mathbb{Z})$-equivalent.

So, basically, whenever we're dealing with $\mathrm{SL}_2(\mathbb{Z})$-invariance of some kind, we can restrict our analysis to F. A side note: The generating matrices S and T transform F in predictable and visually lovely ways (see Figure 2.2). We now define modular functions.

Definition 2.3.1. *A function $f : \mathbb{H} \to \mathbb{C}$ is called a* **modular function** *if it satisfies the following:*

1. *f is meromorphic in the upper half-plane $\mathbb{H}$.*
2. *$f(\gamma \cdot \tau) = \tau$ for any $\gamma \in SL_2(\mathbb{Z})$ and all $\tau \in \mathbb{H}$.*
3. *The function f has a Fourier expansion at $i\infty$ of the form*

$$f(\tau) = \sum_{n=-m}^{\infty} a(n) q^n,$$

where $q = e^{2\pi i \tau}$, and m is some integer.

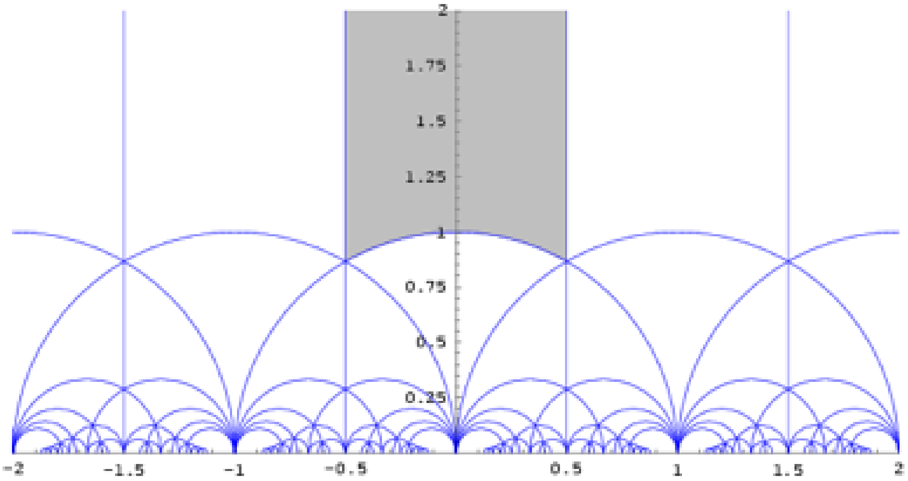


Figure 2.2: Each region bounded by the circles and lines of this figure represent a fundamental region for Γ, reached by transforming F via some combination of matrices S and T. [Arm]

Property 2 gives us modularity; that is, f is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$. Some notes on Property 3: First of all, when considering the behavior of f at infinity, we mean $i\infty$. The reason for this is due to the $\mathrm{SL}_2(\mathbb{Z})$-invariance of f; moving toward z-values with real part approaching ∞, the function values just keep cycling through the same values – picture moving your input points to the left through the Fundamental Domain F.

Secondly, this Fourier expansion in property 3 is called the **q-expansion of** f and can be thought of as the Fourier expansion at $i\infty$, since $\lim_{\tau \to i\infty} e^{2\pi i \tau} = 0$. If $m > 0$, then f has a pole of order m at $i\infty$. If $m \leq 0$, then f is analytic at $i\infty$. Another way of stating property 3 is that f is **meromorphic at the cusp** or **meromorphic at** ∞.

Remark: Modular functions are a specific case of a more general type of function called a modular form. Basically, a modular form is like a modular function but with a slightly modified notion of $\mathrm{SL}_2(\mathbb{Z})$-invariance. More precisely, a **modular form** $f(\tau)$ **of weight** k is a function, either meromorphic or holomorphic on $\mathbb{H}$, that has Property 3 of Definition 2.3.1, and $f(\gamma \cdot \tau) = (c\tau + d)^k f(\tau)$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. In particular, a modular function is a modular form of weight 0.

It turns out that holomorphic modular forms of weight k form a finite dimensional vector space M_k, with a basis given by certain products of G_4 and G_6. Hence we can consider certain linear operators that map M_k to M_k. One such operator, which we will need in a later proof, is called a Hecke operator. The following definition and lemma can be found in [Apo90].

Definition 2.3.2. *For a fixed integer k and any $n \in \mathbb{N}$, the n***th Hecke operator** T_n *is defined on M_k by*

$$(T_n f)(\tau) = n^{k-1} \sum_{d \mid n} d^{-k} \sum_{b=0}^{d-1} f\left(\frac{n\tau + bd}{d^2}\right).$$

What we are really interested in, though, is the Fourier expansion of such Hecke-operated functions:

Lemma 2.3.3. *If $f \in M_k$ and f has the Fourier expansion*

$$f(\tau) = \sum_{m=0}^{\infty} c(m) q^m,$$

then $T_n f$ has the Fourier expansion

$$(T_n f)(\tau) = \sum_{m=0}^{\infty} \left(\sum_{d|(m,n)} d^{k-1} c\left(\frac{mn}{d^2}\right) \right) q^m.$$

Thus if our function f is a modular function; that is, an element of M_0, this expansion becomes

$$(T_n f)(\tau) = \sum_{m=0}^{\infty} \left(\sum_{d|(m,n)} \frac{1}{d} \cdot c\left(\frac{mn}{d^2}\right) \right) q^m .$$

Further exploration of modular functions can be found in [Apo90] and [Ser73].

2.4 Jay at Last!

Now we finally have the tools to define the j-function. Let's do so!

Definition 2.4.1. *The j-invariant of a lattice L is defined as the complex number*

$$j(L) = 1728 \cdot \frac{g_2(L)^3}{g_2(L)^3 - 27 g_3(L)^2},$$

where g_2 and g_3 are the previously-defined modified Eisenstein series.

Note that the denominator of j is $g_2(L)^3 - 27g_3(L)^2 = \Delta(L)$. This is handy, since Lemma 2.2.8 tells us that $\Delta(L)$ is never 0 for any lattice L. The j-function itself is defined like so:

Definition 2.4.2. *The j-**function** $j(\tau)$ is defined as a complex-valued function such that, for any $\tau \in \mathbb{H}$,*

$$j(\tau) = j([1, \tau]),$$

that is, $j(\tau)$ is the j-invariant of the lattice $[1, \tau]$.

Similarly, we modify our old series g_2 and g_3 so that

$$g_2(\tau) = g_2([1, \tau]) = 60 \sum_{m,n} \frac{1}{(m + n\tau)^4}$$

and

$$g_3(\tau) = g_3([1, \tau]) = 140 \sum_{m,n} \frac{1}{(m + n\tau)^6}.$$

So then we have that

$$j(\tau) = 1728 \cdot \frac{g_2(\tau)^3}{g_2(\tau)^3 - 27 g_3(\tau)^2}.$$

The j-function has many important properties, such as the following:

Lemma 2.4.3. *The j-function is holomorphic on the upper half plane $\mathbb{H}$.*

We will use this lemma as a fact. A good proof, relying on the uniform convergence of g_2 and g_3 can be found in [Cox89].

Chapter 3

Properties of $j(\tau)$

3.1 Proof of Theorem 1.2.4

Recall that Theorem 1.2.4 states that $j(\tau)$ is a bijection between $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ and $\mathbb{C}$. Our first step is to show that $j(L)$, as an invariant, characterizes a lattice L with respect to its homothetic lattices. (Recall the definition of *homothetic*, Definition 2.2.1.)

Theorem 3.1.1. *If L and L' are lattices in $\mathbb{C}$, then $j(L) = j(L')$ if and only if L and L' are homothetic.*

Proof. ($\Leftarrow$) Suppose L and L' are homothetic; that is, $L' = \lambda L$ for some $\lambda \in \mathbb{C}$. We know

$$g_2(L') = g_2(\lambda L) = 60 \sum_{\omega \in L\backslash\{0\}} \frac{1}{(\lambda\omega)^4} = \frac{1}{\lambda^4} 60 \sum_{\omega \in L\backslash\{0\}} \frac{1}{\omega^4} = \lambda^{-4} g_2(L).$$

Similarly,

$$g_3(L') = g_3(\lambda L) = 140 \sum_{\omega \in L\backslash\{0\}} \frac{1}{(\lambda\omega)^6} = \frac{1}{\lambda^6} 140 \sum_{\omega \in L\backslash\{0\}} \frac{1}{\omega^6} = \lambda^{-6} g_3(L).$$

So then

$$j(L') = 1728 \cdot \frac{g_2(L')^3}{g_2(L')^3 - 27 g_3(L')^2} = 1728 \cdot \frac{\lambda^{-12} g_2(L)^3}{\lambda^{-12} g_2(L)^3 - 27\lambda^{-12} g_3(L)^2}$$
$$= 1728 \cdot \frac{g_2(L)^3}{g_2(L)^3 - 27 g_3(L)^2} = j(L).$$

□

Now we prove the reverse implication.
($\Rightarrow$) Let $j(L) = j(L')$. Suppose we can find a $\lambda \in \mathbb{C}$ such that $g_2(L') = \lambda^{-4} g_2(L)$ and $g_3(L') = \lambda^{-6} g_3(L)$. Then $g_2(L') = g_2(\lambda L)$ and $g_3(L) = g_3(\lambda L)$, and by Lemma 2.2.6, we have that the Laurent expansion for $\wp(z; L')$ is

$$\wp(z; L') = \frac{1}{z^2} + \sum_{n=1}^{\infty} p(g_2(L'), g_3(L')) z^{2n} = \frac{1}{z^2} + \sum_{n=1}^{\infty} p(g_2(\lambda L), g_3(\lambda L)) z^{2n} = \wp(z; \lambda L).$$

Thus $\wp(z; L')$ and $\wp(z; \lambda L)$ have the same Laurent expansion about 0. So then these two functions agree on a neighborhood U about the origin. But we have that $\wp(z; L')$ and $\wp(z; \lambda L)$ are analytic on the region $\Omega := \mathbb{C} \setminus (\lambda L \cup L')$, and the set

$$\{z \in \Omega : \wp(z; L') = \wp(z; \lambda L)\}$$

certainly has a limit point in $U \cap \Omega$, and hence

$$\wp(z; L') = \wp(z; \lambda L)$$

on all of Ω, so $\wp(z; L')$ and $\wp(z; \lambda L)$ must have the same poles. Since, by Lemma 2.2.5, the lattice L' is precisely the set of poles of $\wp(z; L')$, then $L' = \lambda L$. Thus, to complete this proof, we need only find a $\lambda \in \mathbb{C}$ such that $g_2(L') = \lambda^{-4} g_2(L)$ and $g_3(L') = \lambda^{-6} g_3(L)$.

Note that

$$g_2(L)^3 - 27 g_3(L)^2 = \Delta(L)$$

can never be 0 by Lemma 2.2.8. Thus, we have the following three cases:

- Case 1: $g_2(L') = 0$ and $g_3(L') \neq 0$. Then choose λ so that $\lambda^6 = \frac{g_3(L)}{g_3(L')}$.
- Case 2: $g_3(L') = 0$ and $g_2(L') \neq 0$. Then choose λ so that $\lambda^4 = \frac{g_2(L)}{g_2(L')}$.
- Case 3: $g_2(L') \neq 0$ and $g_3(L') \neq 0$. Then choose λ such that $\lambda^4 = \frac{g_2(L)}{g_2(L')}$. Since we have that $j(L) = j(L')$, then

$$1728 \cdot \frac{g_2(L)^3}{g_2(L)^3 - 27 g_3(L)^2} = 1728 \cdot \frac{g_2(L')^3}{g_2(L')^3 - 27 g_3(L')^2}.$$

By substituting $\lambda^4 g_2(L')$ for $g_2(L)$, we get that

$$\frac{(\lambda^4 g_2(L'))^3}{(\lambda^4 g_2(L'))^3 - 27 g_3(L)^2} = \frac{g_2(L')^3}{g_2(L')^3 - 27 g_3(L')^2}.$$

Cross-multiplying and solving for λ^{12} yields that

$$\lambda^{12} = \frac{-27 g_3(L)^2}{-27 g_3(L')^2} = \frac{g_3(L)^2}{g_3(L')^2}.$$

So then

$$\lambda^6 = \pm \frac{g_3(L)}{g_3(L')}.$$

Assume the sign on the right is +, since if not, we can replace λ by $i\lambda$. Thus we have that $g_3(L') = \lambda^{-6} g_3(L)$.

Hence, in any of the cases, we can find a $\lambda \in \mathbb{C}$ such that $g_2(L') = \lambda^{-4} g_2(L)$ and $g_3(L') = \lambda^{-6} g_3(L)$.

□

Remark: We just showed that the j-invariant characterizes lattice homothety classes. Now, given some lattice $L = [\omega_1, \omega_2]$, there is a homothetic lattice $L' = [1, \frac{\omega_2}{\omega_1}]$ (by letting $\lambda = \frac{1}{\omega_1}$, renumbering if necessary so that $\frac{\omega_2}{\omega_1} \in \mathbb{H}$). But then, by Theorem 3.1.1 and Definition 2.4.2,

$$j(L) = j(L') = j\left(\frac{\omega_2}{\omega_1}\right)$$

. Hence the j-function in a sense represents all possible lattice homothety classes.

Theorem 3.1.1 leads directly to a very nice property of the j-function: invariance under the action of $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 3.1.2. *If $\tau, \tau' \in \mathfrak{h}$, then $j(\tau) = j(\tau')$ if and only if $\tau' = \gamma\tau$ for some $\gamma \in SL_2(\mathbb{Z})$. In particular, $j(\tau)$ is $SL_2(\mathbb{Z})$-invariant.*

Proof. ($\Rightarrow$) Suppose $j(\tau) = j(\tau')$. By definition of the j-function, and by Theorem 3.1.1, then the lattices $L := [1, \tau]$ and $L' := [1, \tau']$ are homothetic. That is, $[1, \tau'] = [\lambda, \lambda\tau]$ for some $\lambda \in \mathbb{C}$.

Hence, we can write

$$\lambda = r\tau' + s \text{ and } \lambda\tau = p\tau' + q,$$

with $p, q, r, s \in \mathbb{Z}$. We then have that

$$\tau = \frac{\lambda\tau}{\lambda} = \frac{p\tau' + q}{r\tau' + s} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \cdot \tau'.$$

Thus we just need to show that $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$.

We have that

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} \tau' \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda\tau \\ \lambda \end{pmatrix}.$$

Similarly, since $[1, \tau'] = [\lambda, \lambda\tau]$, there exists a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with integer entries such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda\tau \\ \lambda \end{pmatrix} = \begin{pmatrix} \tau' \\ 1 \end{pmatrix}.$$

But then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} \tau' \\ 1 \end{pmatrix} = \begin{pmatrix} \tau' \\ 1 \end{pmatrix}.$$

Define the matrix

$$A := \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

So $A \left(\begin{smallmatrix} \tau' \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \tau' \\ 1 \end{smallmatrix}\right)$. Hence $a'\tau' + b' = \tau'$. Divide through by τ' and we get that $a' + b'\frac{1}{\tau'} = 1$. But $\frac{1}{\tau'} \notin \mathbb{R}$, since $[1, \tau']$ is a lattice. Hence, since a' and b' are integers, $b' = 0$ and $a' = 1$. Similarly, we have that $c'\tau' + d' = 1$, leading to the conclusion that $c' = 0$ and $d' = 1$. Thus we have that $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, and so

$$\det(A) = 1 = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \det \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

Since the determinants on the right must be integers, then $\det \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = \pm 1$. So now we just have to show that $\det \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) > 0$.

Let $\tau' = a + bi$ for $a, b \in \mathbb{R}$. Then

$$\mathrm{Im}(\tau) = \mathrm{Im}\left(\frac{p\tau' + q}{r\tau' + s}\right) = \frac{b(ps - qr)}{|r\tau + s|^2} = \frac{\mathrm{Im}(\tau')(ps - qr)}{|r\tau + s|^2}.$$

Since τ and τ' are in $\mathbb{H}$, then $ps - qr = \det\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) > 0$, so

$$\det\begin{pmatrix} p & q \\ r & s \end{pmatrix} = 1,$$

and we have that $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$.

Now we prove the other direction of the implication, which proves to be a bit more straightforward:

($\Leftarrow$) Suppose $\tau' = \gamma\tau$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. That is,

$$\tau' = \frac{p\tau + q}{r\tau + s},$$

with $p, q, r, s \in \mathbb{Z}$, and $\det\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = 1$.

Let $\lambda = r\tau + s$. Then we have that

$$\lambda[1, \tau'] = (r\tau + s)\left[1, \frac{p\tau + q}{r\tau + s}\right] = [r\tau + s, p\tau + q].$$

But we can write

$$-q(r\tau + s) + s(p\tau + q) = (ps - qr)\tau + (qs - qs) = \tau,$$

since $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Similarly,

$$p(r\tau + s) - r(p\tau + q) = (pr - pr)\tau + (ps - qr) = 1.$$

Thus $[1, \tau] \subseteq \lambda[1, \tau']$, and since clearly $\lambda[1, \tau'] \subseteq [1, \tau]$, we have that

$$[r\tau + s, p\tau + q] = \lambda[1, \tau'] = [1, \tau],$$

and hence $[1, \tau']$ is homothetic to $[1, \tau]$. It follows from Theorem 3.1.1 that $j(\tau) = j(\tau')$. □

Let us now think of Theorem 3.1.2 in the context of Theorem 1.2.4, what we aim to prove. The forward implication of Theorem 3.1.2 gives that $j(\tau)$, viewed as a map from $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ to $\mathbb{C}$, is well-defined. The backwards implication gives us that this map is one-to-one. Hence we already have half of Theorem 1.2.4 proved; namely, that $j(\tau)$ is one-to-one as a map from $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ to $\mathbb{C}$. Now we wish to show surjectivity. First, though, we must examine its behavior at the cusp, $i\infty$.

Lemma 3.1.3. *We have that*

$$\lim_{\mathrm{Im}(\tau)\to\infty} j(\tau) = \infty.$$

Proof. To start, consider

$$g_2(\tau) = 60\sum_{m,n} \frac{1}{(m + n\tau)^4} = 60\left(2\sum_{m=1}^{\infty} \frac{1}{m^4} + \sum_{m,n;n\neq 0} \frac{1}{(m + n\tau)^4}\right). \tag{3.1}$$

Let $\tau = a + bi$. But if we consider a single term in the right-hand sum, we find that

$$\lim_{\operatorname{Im}(\tau)\to\infty} \frac{1}{(m+n\tau)^4} = \lim_{\operatorname{Im}(\tau)\to\infty} \frac{1}{m^4 + 4m^3(n\tau) + 6m^2(n\tau)^2 + 4m(n\tau)^3 + (n\tau)^4}.$$

Because $n \neq 0$, then the b^4 term in $(n\tau)^4 = n^2(a^4 + 4a^3(bi) - 6a^2b^2 + 4a(bi)^3 - b^4)$ dominates the denominator as b becomes arbitrarily large. Hence,

$$\lim_{\operatorname{Im}(\tau)\to\infty} \frac{1}{(m+n\tau)^4} = 0.$$

Since $g_2(\tau)$ is uniformly convergent by Lemma 2.2.3, then, in taking the limit, equation (3.1) becomes

$$\lim_{\operatorname{Im}(\tau)\to\infty} g_2(\tau) = 120 \sum_{m=1}^{\infty} \frac{1}{m^4}.$$

But this is a known infinite sum, with $\sum_{m=1}^{\infty} \frac{1}{m^4} = \frac{\pi^4}{90}$. Hence,

$$\lim_{\operatorname{Im}(\tau)\to\infty} g_2(\tau) = \frac{4}{3}\pi^4.$$

The limit behavior of $g_3(\tau)$ is shown similarly, with the sum $\sum_{m=1}^{\infty} \frac{1}{m^6} = \frac{\pi^6}{945}$ giving us that

$$\lim_{\operatorname{Im}(\tau)\to\infty} g_3(\tau) = \frac{8}{27}\pi^6.$$

Combining these two results, we get the limit of the denominator of the j-function:

$$\lim_{\operatorname{Im}(\tau)\to\infty} \left[g_2(\tau)^3 - 27g_3(\tau)^2\right] = \left(\frac{4}{3}\pi^4\right)^3 - 27\left(\frac{8}{27}\pi^6\right)^2 = 0. \tag{3.2}$$

Thus, it follows that

$$\lim_{\operatorname{Im}(\tau)\to\infty} j(\tau) = \infty.$$

□

Remark: Since, as shown in equation (3.2), $\Delta(\tau)$ has a simple zero at ∞, then $j(\tau)$ has a simple pole at ∞, and hence $j(\tau)$ is meromorphic at ∞. Combining this fact with Theorem 3.1.2, we have the following immediate corollary:

Corollary 3.1.4. *The function $j(\tau)$ is a modular function for $SL_2(\mathbb{Z})$.*

Now we can prove that $j(\tau)$ is, indeed, surjective.

Theorem 3.1.5. *The function $j : \mathbb{H} \to \mathbb{C}$ is surjective.*

Proof. We know that $j(\tau)$ is nonconstant on $\mathbb{H}$, since there are values for τ that are distinct under the action of $\mathrm{SL}_2(\mathbb{Z})$ (each point in the fundamental region F is $\mathrm{SL}_2(\mathbb{Z})$-distinct from every other point in F). Also, by Lemma 2.4.3, $j(\tau)$ is holomorphic on $\mathbb{H}$. Hence, by the open mapping theorem, the image of $j(\tau)$ must be an open set in $\mathbb{C}$. Since the only set that is both open and closed in $\mathbb{C}$ is itself, it is sufficient to prove that $j(\mathbb{H})$ is closed.

Let $j(\tau_k)$ be a sequence in $j(\mathbb{H})$ converging to some $w \in \mathbb{C}$ (where $\tau_k \in \mathbb{H}$). Since $j(\tau)$ is invariant under $\mathrm{SL}_2(\mathbb{Z})$, then by only considering τ in our fundamental domain F, we may assume that each τ_k is such that

$$|\mathrm{Re}(\tau_k)| \leq \frac{1}{2} \ \text{ and } \ |\mathrm{Im}(\tau_k)| \geq \frac{\sqrt{3}}{2}.$$

Suppose the imaginary parts of the τ_k's are unbounded. But then, by Lemma 3.1.3, $j(\tau_k)$ contains a subsequence converging to ∞. Since $j(\tau_k)$ converges to w, this cannot happen. So the imaginary parts of the τ_k's are bounded, say by some $M \in \mathbb{R}$. Hence each τ_k lies in the region

$$R = \left\{ \tau \in \mathbb{H} : |\mathrm{Re}(\tau)| \leq \frac{1}{2}, \frac{\sqrt{3}}{2} \leq |\mathrm{Im}(\tau)| \leq \mathrm{M} \right\},$$

a compact subspace of $\mathbb{H}$. But this implies that τ_k has a subsequence converging to some $\tau_0 \in \mathbb{H}$. Since $j(\tau)$ is continuous and $j(\tau_k)$ converges to w, then $j(\tau_0) = w$, and hence $w \in j(\mathbb{H})$. Thus $j(\mathbb{H})$ is closed, and so $j(\mathbb{H}) = \mathbb{C}$.

□

Thus, through combining Theorem 3.1.2 and Theorem 3.1.5, we have proven Theorem 1.2.4, that $j(\tau)$ is a bijection between $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ and $\mathbb{C}$.

3.2 The Function $j(\tau)$ as a "Modular Function Generator"

Now we turn our sights to Theorem 1.2.5, namely showing that $j(\tau)$ generates all modular functions on $\mathrm{SL}_2(\mathbb{Z})$. Before the big results, we prove a quick lemma:

Lemma 3.2.1. *A modular function f which is holomorphic at ∞ (i.e., its q-expansion has no negative powers of q) is constant.*

Proof. It suffices to show that $f(\mathbb{H} \cup \{\infty\})$ is compact in $\mathbb{C}$; the maximum modulus principle of complex analysis then tells us that f must be constant. We do so by showing that $f(\mathbb{H} \cup \{\infty\})$ is sequentially compact, i.e., every sequence has a subsequence that converges to a point in $f(\mathbb{H} \cup \{\infty\})$. Since $\mathbb{C}$ is a metric space, compactness is equivalent to sequential compactness.

Let $\{f(\tau_k)\}$ be a sequence in $f(\mathbb{H} \cup \{\infty\})$. Since f is modular and thus $\mathrm{SL}_2(\mathbb{Z})$-invariant, we can assume that each τ_k lies in our fundamental region F. If the imaginary parts of the τ_k's are unbounded, then there is a subsequence of $\{\tau_k\}$ converging to $i\infty$. But then there is a subsequence of $\{f(\tau_k)\}$ which converges to $f(\infty)$, which is a finite complex number since f is holomorphic at ∞.

If the imaginary parts of the τ_k's are bounded, say $\mathrm{Im}(\tau_k) \leq M$ for some $M \in \mathbb{R}$, then each τ_k lies in the region $R := \{z \in \mathbb{C} : |z| \geq 1, -1/2 \leq \mathrm{Re}(z) \leq 1/2, \mathrm{Im}(z) \leq M\}$, which is closed and bounded and hence compact. Thus a subsequence of $\{\tau_k\}$ can be found which converges to an element τ_0 of R, and hence $f(\tau_k)$ converges to $f(\tau_0)$, since f is continuous. Hence, $f(\mathbb{H} \cup \{\infty\})$ is compact and thus f must be constant. □

Lemma 3.2.2. *Every holomorphic modular function for $SL_2(\mathbb{Z})$ is a polynomial in $j(\tau)$.*

Proof. Suppose $f(\tau)$ is a holomorphic modular function for $\mathrm{SL}_2(\mathbb{Z})$. Since f is modular and thus meromorphic at the cusp $i\infty$, its q-expansion looks like

$$f(\tau) = \sum_{n=-m}^{\infty} a_n q^n,$$

where m is some positive integer. But the only negative q-power term in $j(\tau)$'s q-expansion is simply q^{-1}. Hence, we can define a polynomial $A(x)$ such that $f(\tau) - A(j(\tau))$ has no terms with negative q powers. So then $f(\tau) - A(j(\tau))$ is holomorphic at ∞ and hence must be constant by Lemma 3.2.1. Thus,

$$f(\tau) = k + A(j(\tau))$$

for some $k \in \mathbb{C}$ and hence $f(\tau)$ is a polynomial in $j(\tau)$. □

And now our goal for this section:

Theorem 3.2.3. *Every modular function for $SL_2(\mathbb{Z})$ is a rational function in $j(\tau)$.*

Proof. Suppose $f(\tau)$ is a modular function for $\mathrm{SL}_2(\mathbb{Z})$. Our goal is to find some polynomial $B(x)$ such that $B(j(\tau))f(\tau)$ is holomorphic. Then applying Lemma 3.2.2, we obtain the desired conclusion.

Because f is modular, it only has a finite number of poles in our fundamental domain F. The idea is to find a polynomial of $j(\tau)$ for each pole that kills each pole of $f(\tau)$, making the resulting function holomorphic at that point.

Let τ_0 be a pole of f of order m. Suppose $j'(\tau_0) \neq 0$. Then consider

$$(j(\tau) - j(\tau_0))^m f(\tau).$$

If we write $j(\tau)$ and $f(\tau)$ as Laurent series about τ_0, we get

$$j(\tau) = \sum_{n=0}^{\infty} a_n(\tau - \tau_0)^n \text{ and } f(\tau) = \sum_{n=-m}^{\infty} b_n(\tau - \tau_0)^n,$$

for some constants a_n and b_n since, at the point τ_0, we know j is holomorphic and f has a pole of order m. Note that $a_0 = j(\tau_0)$.

Then consider

$$(j(\tau) - j(\tau_0))^m f(\tau) = \left(\left(\sum_{n=0}^{\infty} a_n(\tau - \tau_0)^n \right) - a_0 \right)^m f(\tau)$$

$$= \left(\sum_{n=1}^{\infty} a_n(\tau - \tau_0)^n \right)^m \sum_{n=-m}^{\infty} b_n(\tau - \tau_0)^n.$$

The resulting expansion will have no negative powers of $(\tau - \tau_0)$, and hence

$$(j(\tau) - j(\tau_0))^m f(\tau)$$

is holomorphic at τ_0.

Multiplying all such polynomials $(j(\tau) - j(\tau_k))^{m_k}$ corresponding to each of the k poles of f in F yields the polynomial in $j(\tau)$

$$\prod_k (j(\tau) - j(\tau_k))^{m_k}$$

which, when multiplied by f, gives a function that is holomorphic at each pole τ_k.

The only case we have omitted is if we have a pole τ_0 such that $j'(\tau_0) = 0$. It turns out that this only happens at i and $e^{2\pi i/3}$, and similar polynomials in $j(\tau)$ are easily obtained at each point so that their product with $f(\tau)$ gives a suitably holomorphic function. [Cox89]

Thus, multiplying all of our collected polynomials together to get some $B(j(\tau))$, we have that $B(j(\tau))f(\tau)$ is holomorphic, and hence by our previous theorem,

$$f(\tau) = \frac{A(j(\tau))}{B(j(\tau))}$$

for some polynomial $A(x)$.

□

Hence, in proving Theorem 1.2.5, we have shown that $j(\tau)$ can be thought of as the generator of all meromorphic modular functions. Such a function in general is called a **hauptmodul**, a term that will crop up in the statement of monstrous moonshine.

3.3 Denominator Formula

The connection between the j-function and the Monster that we alluded to in the Introduction is actually a specific case of the monstrous moonshine theorem. In fact, Borcherds uses yet another property of $j(\tau)$ to complete a portion of his argument in the proof of moonshine. This property is related to something called the denominator formula [Bor92], and takes the following form:

Theorem 3.3.1. *[Bor92] For $p = e^{2\pi i z}$ and $q = e^{2\pi i \tau}$ with $z, \tau \in \mathbb{H}$, we have that*

$$p^{-1} \prod_{m>0, n\in\mathbb{Z}} (1 - p^m q^n)^{c(mn)} = j(p) - j(q) \tag{3.3}$$

where $j(q) - 744 = \sum_{n=-1}^{\infty} c(n)q^n = q^{-1} + 196884q + \cdots$, and $j(p) := j(z)$ and $j(q) := j(\tau)$.

Proof. We begin by multiplying the left-hand side of equation (3.3) by p and then taking the natural logarithm. So we have

$$\ln\left(\prod_{m>0, n\in\mathbb{Z}} (1 - p^m q^n)^{c(mn)}\right) = \sum_{m>0}\sum_{n\in\mathbb{Z}} \ln(1 - p^m q^n)^{c(mn)} = \sum_{m>0}\sum_{n\in\mathbb{Z}} c(mn)\ln(1 - p^m q^n)$$

by basic logarithm rules. Since $p^m q^n = e^{2\pi i(mz+n\tau)}$ and $c(mn) = 0$ for $n < -1$, then we have that

$$0 < |p^m q^n| < 1.$$

Hence we can use the Laurent expansion for the logarithm in our expression above, and we get

$$\sum_{m>0} \sum_{n\in\mathbb{Z}} \sum_{k>0} -c(mn)\frac{(p^m q^n)^k}{k} = -\sum_{m>0} \sum_{n\in\mathbb{Z}} \sum_{k>0} c(mn)\frac{p^{mk} q^{nk}}{k}.$$

Set $m_0 = mk$, and $n_0 = nk$. Then we have

$$-\sum_{m>0} \sum_{n\in\mathbb{Z}} \sum_{k>0} c(mn)\frac{p^{m_0} q^{n_0}}{k}.$$

By reindexing, and considering the sum over particular m_0's and n_0's instead of k's, we get

$$-\sum_{m>0} \sum_{n\in\mathbb{Z}} \sum_{0<k|(m,n)} \frac{1}{k} c\left(\frac{mn}{k^2}\right) p^m q^n. \tag{3.4}$$

But, by Lemma 2.3.3, we know that, for fixed $m > 0$,

$$\sum_{0<k|(m,n)} \sum_{n\in\mathbb{Z}} \frac{1}{k} c\left(\frac{mn}{k^2}\right) p^m q^n = T_m\left(\sum_{n\in\mathbb{Z}} c(n)q^n\right) p^m,$$

where T_m is the m-th Hecke operator. Hence, expression (3.4) now becomes

$$-\sum_{m>0} T_m\left(\sum_{n\in\mathbb{Z}} c(n)q^n\right) p^m = -\sum_{m>0} T_m(j(q) - 744)p^m.$$

As each T_m is a linear operator on the space of holomorphic modular functions, then this expression can be written

$$\sum_{m>0} f_m(q)p^m,$$

where each f_m is a holomorphic modular function. Now recall that we began by multiplying the left hand side of equation 3.3 by p and taking the logarithm. Hence we have shown that

$$\ln\left(\prod_{m>0, n\in\mathbb{Z}} (1 - p^m q^n)^{c(mn)}\right) = \sum_{m>0} f_m(q)p^m.$$

If we exponentiate our expression and multiply by p^{-1}, then, we get that

$$p^{-1} \prod_{m>0, n\in\mathbb{Z}} (1 - p^m q^n)^{c(mn)} = p^{-1} e^{\sum_{m>0} f_m(q)p^m}.$$

Since e^z is an analytic function with a nice Fourier expansion with respect to p, then this expression becomes

$$p^{-1} \sum_{m\geq 0} h_m(q)p^m,$$

where each h_m is some holomorphic modular function. By Lemma 3.2.2, each $h_m(q)$ can be written as a polynomial in $j(q)$. Thus we get that the left hand side of equation (3.3) is equal to

$$\sum_{m\geq -1} g_m(j(q))p^m,$$

where each g_m is a polynomial. Now we compare this sum with the right hand side of (3.3),

$$j(p) - j(q) = \frac{1}{p} + 744 + \sum_{m\geq 1} c(m)p^m - j(q),$$

clearly a polynomial in $j(q)$. Hence, we have shown that each side of (3.3) can be written as a polynomial of $j(q)$. But, as Borcherds notes in [Bor92], if we know the coefficients of q_n for $n \leq 0$ in a polynomial of $j(q)$, we can determine the polynomial entirely. So it will suffice to show that, in the left hand expansion of (3.3) the coefficient on $\frac{1}{q}$ is -1, the constant term is $j(p) - 744$, and the coefficients for q^n, $n < -1$ are all 0.

First, note that $c(mn)$ is zero whenever $n \leq 0$ and $m > 1$. Hence, the left hand side of (3.3) becomes

$$p^{-1}(1 - p^1q^{-1})^1 \prod_{m>0,n>0} (1 - p^m q^n)^{c(mn)} = (p^{-1} - q^{-1}) \prod_{m>0,n>0} (1 - p^m q^n)^{c(mn)}$$

$$= p^{-1} \prod_{m>0,n>0} (1 - p^m q^n)^{c(mn)} - q^{-1} \prod_{m>0,n>0} (1 - p^m q^n)^{c(mn)}. \tag{3.5}$$

Since $n \geq 1$ in the product, the only $\frac{1}{q}$ term will come from the right-hand product in expression (3.5), multiplying the only constant term in the product, 1. Hence, the coefficient on $\frac{1}{q}$ is, indeed, -1.

As for terms which are constant in q, in the left-hand product in expression (3.5) we just have p^{-1}. In the right-hand product, the constant terms arise when $n = 1$, since then the q^{-1} distributes through and cancels out the q term. So we wish to find the coefficients on the q terms in the product

$$\prod_{m>0} (1 - p^m q)^{c(m)}. \tag{3.6}$$

By the binomial theorem, $(1 - p^m q)^{c(m)} = 1 - c(m)p^m q + \cdots$ for each $m > 1$. So then expression (3.6) looks like

$$\prod_{m>0} (1 - c(m)p^m q + \cdots),$$

so the only q term in our product occurs when that second term multiplies with all the 1's in the other product terms. Thus if we collect all our q terms together, we get

$$q \sum_{m>0} c(m)p^m = q\left(j(p) - \frac{1}{p} - 744\right).$$

But then the q^{-1} knocks out the q, and the only q-constant term we get from our left hand product in (3.5) is $\frac{1}{p}$, so the constant term in the left-hand side of (3.3) is $j(p) - 744$, the same as the q-constant term in the right-hand side.

Thus, the left and right hand sides of (3.3) are equal as polynomials in $j(q)$ and are hence equal.

$\square$

Chapter 4

Monstrous Moonshine

We at last return to the conjecture that connects this wondrous function j to the wondrous group $\mathbb{M}$. As you might imagine, the result is quite wondrous. First, though, we need just one more object before we can actually state the theorem which Borcherds proved.

4.1 The Moonshine Module $V_\natural$

Now that we've introduced the Monster and Representation Theory, we can briefly introduce the Moonshine Module $V_\natural$. In 1988, Igor Frenkel, James Lepowsky, and Arne Meurman constructed the $\mathbb{M}$-module $V_\natural$ (that is, the monster group $\mathbb{M}$ has a group action on $V_\natural$) [FLM88]. The pivotal property of $V_\natural$ is the fact that it can be considered an infinite-dimensional, graded representation of $\mathbb{M}$. This is not a representation in the strict sense (it is infinite dimensional!); rather, $V_\natural$ can be expressed as a direct sum of representations of the Monster, like so:

$$V_\natural = \bigoplus_{n=-1}^{\infty} V_n,$$

where each V_n is a finite representation of the Monster. The construction of $V_\natural$ takes up much of a mathematical tome of 508 pages [FLM88], and will not explicitly be covered here. Suffice to say, the authors constructed $V_\natural$ such that the respective dimensions of each V_n correspond with the coefficients of the j-function's q-expansion. So this representation helps to shed light on the observation noted in the Introduction. Hence $V_\natural$ was named the Moonshine Module, after Conway and Norton's famous conjecture that this representation helped to prove.

4.2 Once Conjecture, Now Theorem

The monstrous moonshine conjecture of Conway and Norton coalesced over the years since they first posed it, reflecting the new discoveries of various algebraic structures. Eventually the conjecture was transformed into something general enough and clear enough to prove. Here is the actual statement of the theorem which Borcherds proved in [Bor92]:

Theorem 4.2.1. *[Bor92] Let* $V = V_\natural = \bigoplus_{n\in\mathbb{Z}}^{\infty} V_n$ *be the Moonshine Module constructed by Frenkel, Lepowsky, and Meurman. Then, for any* $g \in \mathbb{M}$*, the McKay-Thompson series*

$$T_g(q) = \sum_{n=-1}^{\infty} tr(g|V_n)q^n$$

is a hauptmodul for a genus 0 *subgroup of* $\mathrm{SL}_2(\mathbb{R})$*.*

As mentioned in Chapter 3, a Hauptmodul is, in essence, an analogue to the j-function in that a Hauptmodul generates all modular functions for a particular modular subgroup G of $SL_2(\mathbb{R})$. "Genus 0" describes the topological genus of the resulting surface when we take $\mathbb{H}$ and mod out by the action of our group G, then compactify. For example, with $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, then $\Gamma\backslash\mathbb{H}$ can be viewed as the Riemann sphere missing the point at infinity. For a visualization, look at the fundamental region F. The left-hand side is identified with the right-hand side, so "roll up" one side to the other, then tape the bottom edges together. This gives us the Riemann sphere missing infinity. So if we just add in the point at infinity, we get the Riemann sphere, which is a genus 0 topological space.

How does this theorem tie into the j-function, you ask? Well, if we let $g = 1$, the identity element of the Monster, then because each representation V_n has dimension corresponding to the coefficients of the Fourier expansion of $j(\tau)$, we have that

$$T_1(q) = \frac{1}{q} + 196884q + 2149360q^2 + \cdots = j(\tau) - 744.$$

Of course, we can construct a Thompson series for every element $g \in \mathbb{M}$. Since $\mathbb{M}$ has 194 conjugacy classes, then Theorem 2.1.6 tells us that we can expect at most 194 Thompson-Mckay series T_g. In fact, there are only 171 distinct series [Gan04], and, by the moonshine theorem, every one is a Hauptmodul for some genus 0 subgroup of $\mathrm{SL}_2(\mathbb{R})$.

4.3 The Future

The discovery of this fascinating connection between $\mathbb{M}$ and modular functions has led to a wide range of related research. For instance, the question has been asked, "Are there any similar connections between other groups and modular functions?" In 1987, even before monstrous moonshine was actually proved, Conway and Norton synthesized a generalized moonshine conjecture which essentially explores similar connections while generalizing the characters with respect to V_n (see [Car08]). As it is, this conjecture is still unproven, but many other relevant results have been proved along the way.

The exploration of moonshine has led to the creation and examination of various algebraic objects, including vector operator algebras, Kac-Moody algebras, genus 0 functions, and $V_\natural$ itself. Monstrous moonshine has even found purchase in the realm of physics, where it is used to explore something called conformal field theory, a particular quantum field theory on 2-dimensional space-time (see [Gan04] for a brief overview).

Though moonshine now delves ever deeper into strange, new algebraic objects, none of it would be possible without Jay and his monstrous friend.

Bibliography

[Apo90] Tom M. Apostol. *Modular Functions and Dirichlet Series in Number Theory.* Springer-Verlag New York, Inc., New York, NY, 1990.

[Arm] John Armstrong. Billiards 2. http://unapologetic.wordpress.com/2007/02/12/billiards-2/.

[BB87] J.M. Borwein and P.B. Borwein. *Pi & the AGM: A Study in Analytic Number Theory and Computational Complexity.* Wiley, New York, NY, 1987.

[Bor92] Richard Borcherds. Monstrous moonshine and monstrous lie superalgebras. *Invent. Math.*, 109:405–444, 1992.

[Car08] Scott Carnahan. Generalized moonshine i: Genus zero functions. *arXiv:0812.3440v2 [math.RT]*, 2008.

[CN79] J. H. Conway and S. P. Norton. Monstrous moonshine. *Bull. Lond. Math. Soc.*, 11:308–339, 1979.

[Cox89] David A. Cox. *Primes of the Form* $x^2 + ny^2$. John Wiley & Sons, Inc, Hoboken, NJ, 1989.

[FH91] William Fulton and Joe Harris. *Representation Theory: A First Course.* Springer-Verlag New York, Inc., New York, NY, 1991.

[FLM88] Igor Frenkel, James Lepowsky, and Arne Meurman. *Vertex Operator Algebras and the Monster.* Harcourt Brace Jovanovich, San Diego, CA, 1988.

[Gan04] Terry Gannon. Monstrous moonshine: The first twenty-five years. *arXiv:math/0402345v2 [math.QA]*, 2004.

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic.* Springer-Verlag New York, Inc., New York, NY, 1973.

[Sol01] Ronald Solomon. A brief history of the classification of the finite simple groups. *Bulletin of the American Mathematical Society*, 38(3):315–352, 2001.

arXiv:1004.0956v2 [hep-th] 25 Jun 2010

AEI-2010-048, CALT-68-2783,
IPMU-10-0053, YITP-10-20

April, 2010

Notes on the K3 Surface and the Mathieu group M_{24}

Tohru Eguchi[1], Hirosi Ooguri[2,3,4] and Yuji Tachikawa[5]

[1] *Yukawa Institute for Theoretical Physics,*
Kyoto University, Kyoto 606-8502 Japan

[2] *California Institute of Technology, Pasadena, CA 91125, USA*

[3] *Institute for the Physics and Mathematics of the Universe,*
University of Tokyo, Kashiwa 277-8582, Japan

[4] *Max-Planck-Institut für Gravitationsphysik,*
D-14476 Potsdam, Germany

[5] *School of Natural Sciences, Institute for Advanced Study,*
Princeton, NJ 08540, USA

Abstract

We point out that the elliptic genus of the K3 surface has a natural decomposition in terms of dimensions of irreducible representations of the largest Mathieu group M_{24}. The reason is yet a mystery.

Elliptic genus of a complex D-dimensional hyperKähler manifold M is defined as

$$Z_{ell}(\tau;z)=\mathrm{Tr}_{\mathcal{R}\times\mathcal{R}}(-1)^{F_L+F_R}q^{L_0}\bar{q}^{\bar{L}_0}e^{4\pi izJ^3_{0,L}} \tag{1.1}$$

in terms of the two-dimensional supersymmetric sigma model whose target space is M [1]. As M is assumed to be hyperKähler, the two-dimensional theory has $\mathcal{N}=4$ superconformal algebra as its symmetry. Then L_0 and $\bar{L}_0$ are zero modes of the left- and right-moving Virasoro operators; J^3_0 is the zero mode of the 3rd component of the affine $SU(2)$ algebra; F_L and F_R are the left- and right-moving fermion numbers. Trace is taken over the Ramond sector of the theory. This elliptic genus is a Jacobi form of weight=0 and index $\frac{D}{2}$.

Elliptic genus for K3 surface was explicitly calculated in [2] and is given by

$$Z_{ell}(K3)(\tau;z)=8\left[\left(\frac{\theta_2(\tau;z)}{\theta_2(\tau;0)}\right)^2+\left(\frac{\theta_3(\tau;z)}{\theta_3(\tau;0)}\right)^2+\left(\frac{\theta_4(\tau;z)}{\theta_4(\tau;0)}\right)^2\right]. \tag{1.2}$$

Here $\theta_i(\tau;z)\,(i=2,3,4)$ are Jacobi theta functions. Actually the space of Jacobi forms of weight=0 and index=1 is known to be one-dimensional and thus the above result could have been guessed without explicit computation. We find that $Z_{ell}(K3)(\tau;z=0)=24$ and $Z_{ell}(K3)(\tau;z=1/2)=16+\mathcal{O}(q)$ and thus (1.2) reproduces the Euler number and signature of K3.

In Ref. [2] and more recently in [3] the expansion of the K3 elliptic genus in terms of irreducible representations of $\mathcal{N}=4$ superconformal algebra has been discussed in detail. We first provide the data of representation theory [4, 5]. For a rigorous mathematical exposition, see e.g. [6, 7].

Let us introduce the character formula of the BPS (short) representation of spin $\ell=0$ in Ramond sector with $(-1)^F$ insertion

$$\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z)=\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}\mu(\tau;z), \tag{1.3}$$

$$\mu(\tau;z)=\frac{-ie^{\pi iz}}{\theta_1(\tau;z)}\sum_{n=-\infty}^{\infty}(-1)^n\frac{q^{\frac{1}{2}n(n+1)}e^{2\pi inz}}{1-q^ne^{2\pi iz}}. \tag{1.4}$$

The BPS representation has a non-vanishing index

$$\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z=0)=1. \tag{1.5}$$

We also introduce the character of a non-BPS (long) representation with conformal dimension h

$$q^{h-\frac{3}{8}}\,\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}. \tag{1.6}$$

Then the elliptic genus is expanded as

$$Z_{ell}(K3)(\tau;z)=24\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z)+\Sigma(\tau)\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3} \tag{1.7}$$

where the expansion function $\Sigma(\tau)$ is given by

$$\begin{aligned}\Sigma(\tau) &= -8\left[\mu(\tau;z=\frac{1}{2})+\mu(\tau;z=\frac{1+\tau}{2})+\mu(\tau;z=\frac{\tau}{2})\right] && (1.8)\\ &= -2\,q^{-1/8}\,(1-\sum_{n=1}^{\infty}A_nq^n) && (1.9)\end{aligned}$$

If one uses the relation that the non-BPS representation splits into a sum of BPS representations at the unitarity bound $h=1/4$,

$$q^{-\frac{1}{8}}\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}=2\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z)+\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=\frac{1}{2}}(\tau;z), \tag{1.10}$$

the polar term in Σ may be eliminated and the above decomposition (1.7) can also be written as

$$\begin{aligned}Z_{ell}(K3)(\tau;z) &= 20\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z)-2\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=\frac{1}{2}}(\tau;z)\\ &\quad +2\sum_{n=1}^{\infty}A_nq^{n-\frac{1}{8}}\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}. \qquad (1.11)\end{aligned}$$

The coefficients A_n have been computed explicitly for lower orders by expanding the series (1.8),

n	1	2	3	4	5	6	7	8	9	$\cdots$
A_n	45	231	770	2277	5796	13915	30843	65550	132825	$\cdots$

(1.12)

and it was conjectured that they are all positive integers [8].

On the other hand the asymptotic behavior of A_n at large n has recently been derived using the analogue of the Rademacher expansion of modular forms [3]

$$A_n\approx\frac{2}{\sqrt{8n-1}}e^{2\pi\sqrt{\frac{1}{2}(n-\frac{1}{8})}}. \tag{1.13}$$

It turns out that the above formula (1.13) gives a good estimate of A_n even at smaller values of n and this confirms the positivity of the coefficients A_n. Note that the series $\mu(\tau; z)$ (1.4) has the form of a Lerch sum (or mock theta function) and thus has a complex modular transformation law which involves Mordell's integral. In such a situation we can use the method recently developed by mathematicians [9, 10, 11] and construct the Poincaré-Maas series to derive the above asymptotic formula.

The above table contains a surprise: the first 5 coefficients, $A_1, ..., A_5$, are equal to dimensions of *irreducible* representations of M_{24}, the largest Mathieu group, see Appendix A. The coefficients A_6 and A_7 can also be nicely decomposed as

$$A_6 = 3520 + 10395, \tag{1.14}$$

$$A_7 = 10395 + 5796 + 5544 + 5313 + 2024 + 1771 \tag{1.15}$$

into the sum of dimensions[1]. For $n \geq 8$, it is still possible to decompose A_n into a sum of dimensions of irreducible representations of M_{24}, but decompositions are not as unique.[2]

This observation can be compared to the famous observation of McKay and Thompson[14]: there, the first few terms of the expansion coefficients of $J(q)$,

$$J(q) = \frac{1}{q} + 196884 + 21493760q^2 + \ldots \tag{1.16}$$

could be naturally decomposed into the sum of the dimension of the irreducible representation of the monster simple group. Conway and Norton [15] formulated it in terms of an infinite dimensional graded representation of the monster group $\bigoplus_i V_i$ such that $\dim V_i$ is the coefficient of q^i of $J(q)$, and called this observation the monstrous moonshine. Frenkel, Leopwski and Meurman then found [16] that this representation is naturally assocaited to the two-dimensional string propagating on $\mathbb{R}^{26}/\Lambda/\mathbb{Z}_2$ where Λ is the Leech lattice. See e.g. [17] for a recent review.

---

[1] The tentative decomposition of A_7 shown in the previous of this paper was not correct in view of the later study of twisted elliptic genus in [12, 13]. Here it is corrected according to their papers.

[2] It may also be interesting to point out that $2, 3, 5, 7, 11,$ and 23 appear in prime factorization of A_n more frequently than any other prime numbers and with certain periodicities in n. These are also prime factors of the order of M_{24}.

In our case the existence of a natural vector space whose graded dimension gives $\Sigma(q)$ is guaranteed by construction: it is the Hilbert space of the two-dimensional supersymmetric conformal field theory whose target space is $K3$. The problem is to identify the action of M_{24} on it.[3]

The non-Abelian symplectic symmetry of K3 was studied mathematically by [19, 20]. Mukai enumerated eleven K3 surfaces which possess finite non-Abelian automorphism groups. It turns out that all these groups are various subgroups of the Mathieu group M_{24}, see Appendix B for more details. Is it possible that these automorphism groups at isolated points in the moduli space of K3 surface are enhanced to M_{24} over the whole of moduli space when we consider the elliptic genus? This question is currently under study using Gepner models and matrix factorization.

As discussed in [21], expansion coefficients of elliptic genera of hyperKähler manifolds in general have an exponential growth and are closely related to the black hole entropy. In particular in the case of k-th symmetric product of K3 surfaces we obtain the leading behavior

$$A_n \approx e^{2\pi\sqrt{\frac{k^2}{k+1}n-\left(\frac{k}{2(k+1)}\right)^2}} \tag{1.17}$$

which gives the entropy of the standard D1-D5 black hole $S \approx 2\pi\sqrt{kn}$ at large k ($k = Q_1Q_5$ where Q_1 and Q_5 are the numbers of D1 and D5 branes). Thus the elliptic genus of K3 surface may be considered as describing the multiplicity of microstates of a small black hole with $Q_1 = Q_5 = 1$.

Here the situation is somewhat similar to a model of black hole described by Witten in [22], where microstates of a small black hole span the representation space of the monster group. Although the partition function of the theory is discussed, the relevant CFT is modular invariant separately in left and right sectors and the discussion is effectively the same as considering the elliptic genus.

It will be extremely interesting to see if the Mathieu group M_{24} in fact acts on the elliptic genus of K3.

---

[3] Dong and Mason pursued the analogue of the monstrous moonshine in the case of M_{24}, see e.g. [18] and references therein. So far there is no direct connection of their work and the geometry of K3.

Acknowledgment

The authors thank the hospitality of the Aspen Center of Physics during the workshop "Unity of String Theory," when the observation reported in this paper was made. We would like to thank S. Mukai for discussions.

T. E. is supported in part by Grant in Aid from the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT). H. O. is supported in part by U.S. Department of Energy grant DE-FG03-92-ER40701, the World Premier International Research Center Initiative and a Grant-in-Aid for Scientific Research (C) 20540256 of MEXT, and the Humboldt Research Award. Y. T. is supported in part by the NSF grant PHY-0503584, and by the Marvin L. Goldberger membership at the Institute for Advanced Study.

A Data of M_{24}

The largest of the Mathieu group, M_{24}, has

$$2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 = 244823040 \tag{A.1}$$

elements. There are 26 conjugacy classes and 26 irreducible representations. The character table is given in Table 1, whose data is taken from [23, 24]. The conjugacy class is labeled according to the convention of [24]. In the character table, $e_p^\pm$ stands for

$$e_p^\pm = (\pm\sqrt{-p} - 1)/2. \tag{A.2}$$

The dimensions of the irreducible representations are, in the increasing order,

$$\begin{aligned}&1, 23, 45, 45, 231, 231, 252, 253, 483, 770, 770,\\ &990, 990, 1035, 1035, 1035, 1265, 1771, 2024,\\ &2277, 3312, 3520, 5313, 5796, 5544, 10395.\end{aligned} \tag{A.3}$$

Here the irreducible representations of dimensions

$$45, 231, 770, 990, 1035 \tag{A.4}$$

come in complex conjugate pairs. There is in addition an extra real 1035-dimensional irreducible representation.

1A	2A	3A	5A	4B	7A	7B	8A	6A	11A	15A	15B	14A	14B	23A	23B	12B	6B	4C	3B	2B	10A	21A	21B	4A	12A
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	7	5	3	3	2	2	1	1	1	0	0	0	0	0	0	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
252	28	9	2	4	0	0	0	1	−1	−1	−1	0	0	−1	−1	0	0	0	0	12	2	0	0	4	1
253	13	10	3	1	1	1	−1	−2	0	0	0	−1	−1	0	0	1	1	1	1	−11	−1	1	1	−3	0
1771	−21	16	1	−5	0	0	−1	0	0	1	1	0	0	0	0	−1	−1	−1	7	11	1	0	0	3	0
3520	64	10	0	0	−1	−1	0	−2	0	0	0	1	1	1	1	0	0	0	−8	0	0	−1	−1	0	0
45	−3	0	0	1	e_7^+	e_7^-	−1	0	1	0	0	$-e_7^+$	$-e_7^-$	−1	−1	1	−1	1	3	5	0	e_7^-	e_7^+	−3	0
45	−3	0	0	1	e_7^-	e_7^+	−1	0	1	0	0	$-e_7^-$	$-e_7^+$	−1	−1	1	−1	1	3	5	0	e_7^+	e_7^-	−3	0
990	−18	0	0	2	e_7^+	e_7^-	0	0	0	0	0	e_7^+	e_7^-	1	1	1	−1	−2	3	−10	0	e_7^-	e_7^+	6	0
990	−18	0	0	2	e_7^-	e_7^+	0	0	0	0	0	e_7^-	e_7^+	1	1	1	−1	−2	3	−10	0	e_7^+	e_7^-	6	0
1035	−21	0	0	3	$2e_7^+$	$2e_7^-$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^-$	$-e_7^+$	3	0
1035	−21	0	0	3	$2e_7^-$	$2e_7^+$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^+$	$-e_7^-$	3	0
1035	27	0	0	−1	−1	−1	1	0	1	0	0	−1	−1	0	0	0	2	3	6	35	0	−1	−1	3	0
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^+	e_{15}^-	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^-	e_{15}^+	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^+	e_{23}^-	1	1	−2	−7	10	0	0	0	2	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^-	e_{23}^+	1	1	−2	−7	10	0	0	0	2	−1
483	35	6	−2	3	0	0	−1	2	−1	1	1	0	0	0	0	0	0	3	0	3	−2	0	0	3	0
1265	49	5	0	1	−2	−2	1	1	0	0	0	0	0	0	0	0	0	−3	8	−15	0	1	1	−7	−1
2024	8	−1	−1	0	1	1	0	−1	0	−1	−1	1	1	0	0	0	0	0	8	24	−1	1	1	8	−1
2277	21	0	−3	1	2	2	−1	0	0	0	0	0	0	0	0	0	2	−3	6	−19	1	−1	−1	−3	0
3312	48	0	−3	0	1	1	0	0	1	0	0	−1	−1	0	0	0	−2	0	−6	16	1	1	1	0	0
5313	49	−15	3	−3	0	0	−1	1	0	0	0	0	0	0	0	0	0	−3	0	9	−1	0	0	1	1
5796	−28	−9	1	4	0	0	0	−1	−1	1	1	0	0	0	0	0	0	0	0	36	1	0	0	−4	−1
5544	−56	9	−1	0	0	0	0	1	0	−1	−1	0	0	1	1	0	0	0	0	24	−1	0	0	−8	1
10395	−21	0	0	−1	0	0	1	0	0	0	0	0	0	−1	−1	0	0	3	0	−45	0	0	0	3	0

Table 1: Character table of M_{24}.

B M_{24} and the classical geometry of K3

Here we briefly summarize the relation between the classical geometry of the K3 surface and M_{24}, first found in [19] and elaborated in [20].

Before proceeding, we need to recall the definition of M_{24}. Of many equivalent ways to define it, one that is most understandable to string theorists is to use an even self-dual lattice of dimension 24. Consider the root lattice of A_1 whose generator has squared length 2. Let us denote its weight lattice by A_1^* whose generator has squared length 1/2. Take the 24-dimensional lattice ${A_1}^{24}$. This is even but not self-dual, because the dual lattice is A_1^{*24}. An even self-dual lattice N containing ${A_1}^{24}$ will have the structure

$$ {A_1}^{24} \subset N \subset A_1^{*24}. \tag{B.1} $$

Let $\mathcal{G} = N/{A_1}^{24}$, which is a vector subspace of $A_1^{*24}/{A_1}^{24} \simeq {\mathbb{Z}_2}^{24}$. Let us represent an element of $\mathcal{G}$ by a sequence of twenty-four 0 and 1, and define the weight of a vector to be the number of 1's in it.

The self-duality of N translates to the fact $\mathcal{G}$ is 12-dimensional. The evenness translates to the fact that the weight of every element of $\mathcal{G}$ is a multiple of 4. Let us further demand the vectors of N whose length squared are two are the roots of ${A_1}^{24}$ and not more. Then $\mathcal{G}$ does not have an element with weight 4.

These conditions fix the form of $\mathcal{G}$ uniquely, and $\mathcal{G}$ is known as the extended binary Golay code. M_{24} is defined as the subgroup of the permutation S_{24} of the coordinates of ${\mathbb{Z}_2}^{24}$ which preserves $\mathcal{G}$.

The lattice N thus constructed defines a chiral CFT with $c = 24$ whose current algebra is ${A_1}^{24}$. Therefore M_{24} is the discrete symmetry of this chiral CFT.

Now let us recall that the cohomology lattice of K3,

$$ \Lambda = H^*(K3, \mathbb{Z}) \tag{B.2} $$

is also an even self-dual lattice of dimension 24, but with signature $(4, 20)$. The close connection between M_{24} and the geometry of the K3 surface stems from this fact.

Take a K3 surface S, and let G its symmetry preserving the holomorphic 2-form. Let Λ^G the part of Λ preserved by G, and Λ_G its orthogonal complement. Λ_G is inside the primitive part of $H^{1,1}$, thus it is negative definite. Using Nikulin's result, it can be shown that Λ_G is a sublattice of N. Therefore G is a subgroup of M_{24}.

G cannot be M_{24} itself, however. The action of G on S preserves at least H^0, H^4, $H^{2,0}$, $H^{0,2}$ and the Kähler form. Hence Λ^G is at least five-dimensional, and Λ_G is at most 19-dimensional. This implies that the action of G on N as real linear maps should at least have five-dimensional fixed subspace. This translates to the fact that the action of G on 24 points as a subgroup of M_{24} splits them into at least five orbits.

Similarly, starting from a subgroup G of M_{24} which acts on 24 points with at least five orbits, one can construct the action of G on $H^{1,1}$. Using the global Torelli theorem, this translates to the existence of a K3 surface S whose symmetry is G.

One example is the Fermat quartic, $X^4+Y^4+Z^4+W^4=0$ in $\mathbb{CP}^3$. The symmetry is $(\mathbb{Z}_4)^2 \rtimes S_4$, with 384 elements. This is indeed a subgroup of M_{24} which decomposes 24 points into five orbits, of length 1, 1, 2, 4 and 16.

More examples and details of the analysis can be found in [19] and in [20].

References

[1] E. Witten, "Elliptic Genera and Quantum Field Theory," Commun. Math. Phys. **109** (1987) 525.

[2] T. Eguchi, H. Ooguri, A. Taormina and S. K. Yang, "Superconformal Algebras and String Compactification on Manifolds with $SU(N)$ Holonomy," Nucl. Phys. B **315** 193, 1989.

[3] T. Eguchi and K. Hikami, "Superconformal Algebras and Mock Theta Functions 2. Rademacher Expansion for K3 Surface," Commun. Number Theor. and Phys. **3**, 531–554, 2009 [arXiv:0904.0911 [math-ph]].

[4] T. Eguchi and A. Taormina, "Unitary Representations of $\mathcal{N}=4$ Superconformal Algebra," Phys. Lett. B **196** (1987) 75.

[5] T. Eguchi and A. Taormina, "Character Formulas for the $\mathcal{N}=4$ Superconformal Algebra," Phys. Lett. B **200** (1988) 315.

[6] V. G. Kac, "Vertex algebras for beginners," 2nd ed., University lecture series, vol 10, American Mathematical Society, 1998.

[7] V. G. Kac and M. Wakimoto, "Quantum Reduction and Representation Theory of Superconformal Algebras," Advances in Math. **185** (2004) 400 [arXiv:math-ph/0304011].

[8] H. Ooguri, "Superconformal Symmetry and Geometry of Ricci Flat Kähler Manifolds," Int. J. Mod. Phys. A **4** 4303, 1989; A. Taormina and K. Wendland, unpublished 1989.

[9] S. P. Zwegers, "Mock Theta Functions," Ph.D. Thesis, Universiteit Utrecht, 2002.

[10] K. Bringmann and K. Ono, "The $f(q)$ mock theta function conjecture and partition ranks," Invent. Math. **165** 243-266, 2006; "Coefficients of harmonic Maas forms," preprint 2008.

[11] D. Zagier, "Ramanujan's Mock Theta Functions and Their Applications," Séminaire Bourbaki 60ème année, 2006-2007, n° 986.

[12] M. C. N. Cheng, "K3 Surfaces, $\mathcal{N} = 4$ Dyons, and the Mathieu Group M24," arXiv:1005.5415 [hep-th].

[13] M. R. Gaberdiel, S. Hohenegger and R. Volpato, "Mathieu Twining Characters for K3," arXiv:1006.0221 [hep-th].

[14] J. G. Thompson, "Some Numerology between the Fischer-Griess Monster and the Elliptic Modular Function," Bull. London Math. Soc. **11** (1979) 352.

[15] J. H. Conway and S. P. Norton, "Monstrous Moonshine," Bull. London Math. Soc. **11** (1979) 308.

[16] I. B. Frenkel, J. Lepowsky and A. Meurman, "Vertex Operator Algebras and the Monster," Pure and Applied Math., vol. 134, Academic Press, 1988.

[17] T. Gannon, "Monstrous Moonshine: the First Twenty-Five Years," arXiv:math/0402345.

[18] C. Y. Dong and G. Mason, "An Orbifold Theory of Genus Zero Associated to the Sporadic Group M_{24}," Commun. Math. Phys. **164** (1994) 87.

[19] S. Mukai, "Finite groups of automorphisms of $K3$ surfaces and the Mathieu group," Invent. Math., **94** 183, 1988.

[20] S. Kondo, "Niemeier Lattices, Mathieu Groups and Finite Groups of Symplectic Automorphisms of $K3$ surfaces," Duke Math. Journal, **92** 593, 1998 (with an appendix by S. Mukai).

[21] T. Eguchi and K. Hikami, "N=4 Superconformal Algebra and the Entropy of HyperKähler Manifolds," JHEP **1002** 019, 2010 [arXiv:0909.0410 [hep-th]].

[22] E. Witten, "Three-Dimensional Gravity Revisited," arXiv:0706.3359, 2007.

[23] Mathematical Society of Japan, "Iwanami Suugaku Jiten." 4th Japanese ed., Iwanami Shoten, 2007. (English translation of the 3rd Japanese edition is available as the 2nd English edition of "Encyclopedic Dictionary of Mathematics," MIT press, 1987.)

[24] J. Conway, R. Curtis, S. Norton, R. Parker, and R. Wilson, "Atlas of finite groups. Maximal subgroups and ordinary characters for simple groups." With comput. assist. from J. G. Thackray. Clarendon Press, 1985.

Experimental Mathematics, 20(1):91–96, 2011
Copyright © Taylor & Francis Group, LLC
ISSN: 1058-6458 print
DOI: 10.1080/10586458.2011.544585

Notes on the K3 Surface and the Mathieu Group M_{24}

Tohru Eguchi, Hirosi Ooguri, and Yuji Tachikawa

CONTENTS

1. Introduction and Conclusions
2. Appendix: Data on M_{24}
3. Appendix: M_{24} and the classical geometry of K3
Acknowledgments
References

We point out that the elliptic genus of the K3 surface has a natural decomposition in terms of dimensions of irreducible representations of the largest Mathieu group M_{24}. The reason remains a mystery.

1. INTRODUCTION AND CONCLUSIONS

The elliptic genus of a complex D-dimensional hyper-Kähler manifold M is defined as

$$Z_{\text{ell}}(\tau; z) = \text{Tr}_{\mathcal{R}\times\mathcal{R}}(-1)^{F_L+F_R}\, q^{L_0}\, \bar{q}^{\bar{L}_0}\, e^{4\pi i z J^3_{0,L}}$$

in terms of the two-dimensional supersymmetric sigma model whose target space is M [Witten 87]. Since M is assumed to be hyper-Kähler, the two-dimensional theory has $\mathcal{N} = 4$ superconformal algebra as its symmetry. Then L_0 and $\bar{L}_0$ are zero modes of the left- and right-moving Virasoro operators; J^3_0 is the zero mode of the third component of the affine SU(2) algebra; F_L and F_R are the left- and right-moving fermion numbers. The trace is taken over the Ramond sector of the theory. This elliptic genus is a Jacobi form of weight zero and index $D/2$.

The elliptic genus for the K3 surface was explicitly calculated in [Eguchi et al. 89] and is given by

$$Z_{\text{ell}}(K3)(\tau; z) = 8\left[\left(\frac{\theta_2(\tau;z)}{\theta_2(\tau;0)}\right)^2 + \left(\frac{\theta_3(\tau;z)}{\theta_3(\tau;0)}\right)^2 + \left(\frac{\theta_4(\tau;z)}{\theta_4(\tau;0)}\right)^2\right]. \tag{1–1}$$

Here $\theta_i(\tau; z)\,(i = 2, 3, 4)$ are Jacobi theta functions. Actually, the space of Jacobi forms of weight zero and index one is known to be one-dimensional, and thus the above result could have been guessed without explicit computation. We find that $Z_{\text{ell}}(K3)(\tau; z = 0) = 24$ and $Z_{\text{ell}}(K3)(\tau; z = 1/2) = 16 + \mathcal{O}(q)$, and thus (1–1) reproduces the Euler number and signature of K3.

In [Eguchi et al. 89] and more recently in [Eguchi and Hikami 09], the expansion of the K3 elliptic genus in terms of the irreducible representations of the

2000 AMS Subject Classification: 20D08, 14J28, 58J26
Keywords: K3 surface, elliptic genus, Mathieu groups

$\mathcal{N}=4$ superconformal algebra has been discussed in detail. We first provide the data of representation theory [Eguchi and Taormina 87, Eguchi and Taormina 88]. For a rigorous mathematical exposition, see, for example, [Kac 98, Kac and Wakimoto 04].

Let us introduce the character formula of the BPS (short) representation of spin $\ell = 0$ in the Ramond sector with $(-1)^F$ insertion

$$\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z) = \frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}\,\mu(\tau;z), \tag{1–2}$$

$$\mu(\tau;z) = \frac{-ie^{\pi i z}}{\theta_1(\tau;z)} \sum_{n=-\infty}^{\infty} (-1)^n \frac{q^{\frac{1}{2}n(n+1)}\, e^{2\pi i n z}}{1-q^n e^{2\pi i z}}.$$

The BPS representation has nonvanishing index

$$\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau; z=0) = 1.$$

We also introduce the character of a non-BPS (long) representation with conformal dimension h:

$$q^{h-\frac{3}{8}}\,\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}.$$

Then the elliptic genus is expanded as

$$Z_{\mathrm{ell}}(K3)(\tau;z) = 24\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z) + \Sigma(\tau)\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}, \tag{1–3}$$

where the expansion function $\Sigma(\tau)$ is given by

$$\begin{aligned}\Sigma(\tau) &= -8\left[\mu\left(\tau; z=\frac{1}{2}\right) + \mu\left(\tau; z=\frac{1+\tau}{2}\right)\right. \\ &\qquad \left. + \,\mu\left(\tau; z=\frac{\tau}{2}\right)\right] \\ &= -2\,q^{-1/8}\left(1-\sum_{n=1}^{\infty} A_n q^n\right).\end{aligned} \tag{1–4}$$

If one uses the relation that the non-BPS representation splits into a sum of BPS representations at the unitarity bound $h = 1/4$, then

$$q^{-\frac{1}{8}}\,\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3} = 2\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z) + \mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=\frac{1}{2}}(\tau;z),$$

the polar term in Σ may be eliminated, and the decomposition (1–3) can also be written as

$$\begin{aligned}Z_{\mathrm{ell}}(K3)(\tau;z) &= 20\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=0}(\tau;z) - 2\,\mathrm{ch}^{\hat{R}}_{h=\frac{1}{4},\ell=\frac{1}{2}}(\tau;z) \\ &\quad + 2\sum_{n=1}^{\infty} A_n\, q^{n-\frac{1}{8}}\,\frac{\theta_1(\tau;z)^2}{\eta(\tau)^3}.\end{aligned}$$

The coefficients A_n have been computed explicitly for lower orders by expanding the series (1–4) (see Table 1), and it has been conjectured that they are all positive integers [Ooguri 89, Taormina and Wendlend 89, private communication].

n	1	2	3	4	5	6	7	8	$9\cdots$
A_n	45	231	770	2277	5796	13915	30843	65550	$132825\cdots$

TABLE 1. The initial coefficients A_n of the series (1–4)

On the other hand, the asymptotic behavior of A_n at large n has recently been derived using an analogue of the Rademacher expansion of modular forms [Eguchi and Hikami 09]:

$$A_n \approx \frac{2}{\sqrt{8n-1}}\, e^{2\pi\sqrt{\frac{1}{2}\left(n-\frac{1}{8}\right)}}. \tag{1–5}$$

It turns out that (1–5) gives a good estimate of A_n even at smaller values of n, and this confirms the positivity of the coefficients A_n. Note that the series $\mu(\tau;z)$ (1–2) has the form of a Lerch sum (or mock theta function) and thus has a complex modular transformation law that involves Mordell's integral. In such a situation we can use the method recently developed in [Zwegers 02, Bringmann and Ono 06, Bringmann and Ono 08, Zagier 07] and construct the Poincaré–Maass series to derive the above asymptotic formula.

Table 1 contains a surprise: the first five coefficients, $A_1, \ldots, A_5$, are equal to the dimensions of the *irreducible* representations of M_{24}, the largest Mathieu group; see Section 2. The coefficients A_6 and A_7 can also be nicely decomposed as sums of dimensions:[1]

$$\begin{aligned}A_6 &= 3520 + 10395, \\ A_7 &= 10395 + 5796 + 5544 + 5313 + 2024 + 1771.\end{aligned}$$

For $n \geq 8$, it is still possible to decompose A_n into a sum of dimensions of irreducible representations of M_{24}, but the decompositions are not as unique.[2]

This observation can be compared to the famous observation in [Thompson 79], where the first few terms of the expansion coefficients of $J(q)$,

$$J(q) = \frac{1}{q} + 196884 + 21493760q^2 + \cdots, \tag{1–6}$$

could be naturally decomposed into the sum of the dimension of the irreducible representation of the monster simple group. In [Conway and Norton 79] it is formulated in terms of an infinite-dimensional graded representation of the monster group $\bigoplus_i V_i$ such that $\dim V_i$ is the coefficient of q^i of $J(q)$, and Conway and Norton called this observation monstrous moonshine.

---

[1] The tentative decomposition of A_7 shown in a previous version of this paper was incorrect in view of a later study of twisted elliptic genus in [Cheng 10, Gaberdiel et al. 10]. Here it is corrected according to those papers.

[2] It may also be interesting to point out that $2, 3, 5, 7, 11, 23$ appear in the prime factorization of A_n more frequently than any other prime numbers and with certain periodicities in n. These are also prime factors of the order of M_{24}.

It was then found [Frenkel et al. 88] that this representation is naturally associated with the two-dimensional string propagating on $\mathbb{R}^{26}/\Lambda/\mathbb{Z}_2$, where Λ is the Leech lattice. See, for example, [Gannon 04] for a recent review.

In our case, the existence of a natural vector space whose graded dimension gives $\Sigma(q)$ is guaranteed by the construction: it is the Hilbert space of the two-dimensional supersymmetric conformal field theory whose target space is K3. The problem is to identify the action of M_{24} on it.[3]

The nonabelian symplectic symmetry of K3 was studied mathematically in [Mukai 88, Kondo 98]. Mukai enumerated eleven K3 surfaces that possess finite nonabelian automorphism groups. It turns out that all these groups are various subgroups of the Mathieu group M_{24}; see Section 3 for more details. Is it possible that these automorphism groups at isolated points in the moduli space of the K3 surface are enhanced to M_{24} over the whole moduli space when we consider the elliptic genus? This question is currently under study using Gepner models and matrix factorization.

As discussed in [Eguchi and Hikami 10], expansion coefficients of elliptic genera of hyper-Kähler manifolds in general have an exponential growth and are closely related to the black hole entropy. In particular, in the case of the kth symmetric product of the K3 surfaces we obtain the leading behavior

$$A_n \approx e^{2\pi\sqrt{\frac{k^2}{k+1}n-\left(\frac{k}{2(k+1)}\right)^2}},$$

which gives the entropy of the standard D1-D5 black hole $S \approx 2\pi\sqrt{kn}$ at large k ($k = Q_1Q_5$, where Q_1 and Q_5 are the numbers of D1 and D5 branes). Thus the elliptic genus of the K3 surface may be considered as describing the multiplicity of microstates of a small black hole with $Q_1 = Q_5 = 1$.

Here the situation is somewhat similar to a model of black hole described in [Witten 07], where microstates of a small black hole span the representation space of the monster group. Although the partition function of the theory is discussed, the relevant CFT is modular invariant separately in left and right sectors, and the discussion is effectively the same as considering the elliptic genus.

It will be extremely interesting to see whether the Mathieu group M_{24} in fact acts on the elliptic genus of K3.

2. APPENDIX: DATA ON M_{24}

The largest Mathieu group, M_{24}, has

$$2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 = 244823040$$

elements. There are 26 conjugacy classes and 26 irreducible representations. The character table is given in Table 2, whose data are taken from [Math. Soc. Japan 07, Conway et al. 85]. The conjugacy class is labeled according to the convention of [Conway et al. 85]. In the character table, $e_p^\pm$ stands for

$$e_p^\pm = (\pm\sqrt{-p} - 1)/2.$$

The dimensions of the irreducible representations are, in increasing order,

1, 23, 45, 45, 231, 231, 252, 253, 483, 770, 770, 990, 990, 1035, 1035, 1035, 1265, 1771, 2024, 2277, 3312, 3520, 5313, 5796, 5544, 10395.

Here the irreducible representations of dimensions

$$45,\ 231,\ 770,\ 990,\ 1035$$

come in complex conjugate pairs. There is in addition an extra real 1035-dimensional irreducible representation.

3. APPENDIX: M_{24} AND THE CLASSICAL GEOMETRY OF K3

Here we briefly summarize the relation between the classical geometry of the K3 surface and M_{24}, first found in [Mukai 88] and elaborated in [Kondo 98].

Before proceeding, we need to recall the definition of M_{24}. Of many equivalent ways to define it, one that is most understandable to string theorists is to use an even self-dual lattice of dimension 24. Consider the root lattice of A_1 whose generator has squared length 2. Let us denote its weight lattice by A_1^*, whose generator has squared length 1/2. Take the 24-dimensional lattice ${A_1}^{24}$. This is even but not self-dual, because the dual lattice is A_1^{*24}. An even self-dual lattice N containing ${A_1}^{24}$ will have the structure

$${A_1}^{24} \subset N \subset A_1^{*24}.$$

Let $\mathcal{G} = N/{A_1}^{24}$, which is a vector subspace of $A_1^{*24}/{A_1}^{24} \simeq {\mathbb{Z}_2}^{24}$. Let us represent an element of $\mathcal{G}$ by a

---

[3] Dong and Mason pursued the analogue of monstrous moonshine in the case of M_{24}; see, for example, [Dong and Mason 94] and references therein. So far, there is no direct connection between their work and the geometry of K3.

1A	2A	3A	5A	4B	7A	7B	8A	6A	11A	15A	15B	14A	14B	23A	23B	12B	6B	4C	3B	2B	10A	21A	21B	4A	12A
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	7	5	3	3	2	2	1	1	1	0	0	0	0	0	0	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
252	28	9	2	4	0	0	0	1	−1	−1	−1	0	0	−1	−1	0	0	0	0	12	2	0	0	4	1
253	13	10	3	1	1	1	−1	−2	0	0	0	−1	−1	0	0	1	1	1	1	−11	−1	1	1	−3	0
1771	−21	16	1	−5	0	0	−1	0	0	1	1	0	0	0	0	−1	−1	−1	7	11	1	0	0	3	0
3520	64	10	0	0	−1	−1	0	−2	0	0	0	1	1	1	1	0	0	0	−8	0	0	−1	−1	0	0
45	−3	0	0	1	e_7^+	e_7^-	−1	0	1	0	0	$-e_7^+$	$-e_7^-$	−1	−1	1	−1	1	3	5	0	e_7^-	e_7^+	−3	0
45	−3	0	0	1	e_7^-	e_7^+	−1	0	1	0	0	$-e_7^-$	$-e_7^+$	−1	−1	1	−1	1	3	5	0	e_7^+	e_7^-	−3	0
990	−18	0	0	2	e_7^+	e_7^-	0	0	0	0	0	e_7^+	e_7^-	1	1	1	−1	−2	3	−10	0	e_7^-	e_7^+	6	0
990	−18	0	0	2	e_7^-	e_7^+	0	0	0	0	0	e_7^-	e_7^+	1	1	1	−1	−2	3	−10	0	e_7^+	e_7^-	6	0
1035	−21	0	0	3	$2e_7^+$	$2e_7^-$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^-$	$-e_7^+$	3	0
1035	−21	0	0	3	$2e_7^-$	$2e_7^+$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^+$	$-e_7^-$	3	0
1035	27	0	0	−1	−1	−1	1	0	1	0	0	−1	−1	0	0	0	2	3	6	35	0	−1	−1	3	0
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^+	e_{15}^-	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^-	e_{15}^+	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^+	e_{23}^-	1	1	−2	−7	10	0	0	0	2	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^-	e_{23}^+	1	1	−2	−7	10	0	0	0	2	−1
483	35	6	−2	3	0	0	−1	2	−1	1	1	0	0	0	0	0	0	3	0	3	−2	0	0	3	0
1265	49	5	0	1	−2	−2	1	1	0	0	0	0	0	0	0	0	0	−3	8	−15	0	1	1	−7	−1
2024	8	−1	−1	0	1	1	0	−1	0	−1	−1	1	1	0	0	0	0	0	8	24	−1	1	1	8	−1
2277	21	0	−3	1	2	2	−1	0	0	0	0	0	0	0	0	0	2	−3	6	−19	1	−1	−1	−3	0
3312	48	0	−3	0	1	1	0	0	1	0	0	−1	−1	0	0	0	−2	0	−6	16	1	1	1	0	0
5313	49	−15	3	−3	0	0	−1	1	0	0	0	0	0	0	0	0	0	−3	0	9	−1	0	0	1	1
5796	−28	−9	1	4	0	0	0	−1	−1	1	1	0	0	0	0	0	0	0	0	36	1	0	0	−4	−1
5544	−56	9	−1	0	0	0	0	1	0	−1	−1	0	0	1	1	0	0	0	0	24	−1	0	0	−8	1
10395	−21	0	0	−1	0	0	1	0	0	0	0	0	0	−1	−1	0	0	3	0	−45	0	0	0	3	0

TABLE 2. Character table of M_{24}.

sequence of 24 zeros and ones, and define the weight of a vector to be the number of ones in it.

The self-duality of N translates to the fact that $\mathcal{G}$ is 12-dimensional. The evenness translates to the fact that only the weight of every element of $\mathcal{G}$ is a multiple of 4. Let us further demand that only the vectors of N whose length squared is two be the roots of ${A_1}^{24}$ and not more. Then $\mathcal{G}$ does not have an element with weight 4.

These conditions fix the form of $\mathcal{G}$ uniquely, and $\mathcal{G}$ is known as the extended binary Golay code. The group M_{24} is defined as the subgroup of the permutation group S_{24} of the coordinates of ${\mathbb{Z}_2}^{24}$ that preserves $\mathcal{G}$.

The lattice N thus constructed defines a chiral CFT with $c = 24$ whose current algebra is ${A_1}^{24}$. Therefore M_{24} is the discrete symmetry of this chiral CFT.

Now let us recall that the cohomology lattice of K3,

$$\Lambda = H^*(K3, \mathbb{Z}),$$

is also an even self-dual lattice of dimension 24, but with signature $(4, 20)$. The close connection between M_{24} and the geometry of the K3 surface stems from this fact.

Take a K3 surface S, and let G be its symmetry preserving the holomorphic 2-form. Let Λ^G be the part of Λ preserved by G, and Λ_G its orthogonal complement. Then Λ_G is inside the primitive part of $H^{1,1}$, and thus it is negative definite. Using Nikulin's result, it can be shown that Λ_G is a sublattice of N. Therefore G is a subgroup of M_{24}.

The subgroup G cannot be M_{24} itself, however. The action of G on S preserves at least H^0, H^4, $H^{2,0}$,

$H^{0,2}$, and the Kähler form. Hence Λ^G is at least five-dimensional, and Λ_G is at most 19-dimensional. This implies that the action of G on N as real linear maps should at least have a five-dimensional fixed subspace. This translates to the fact that the action of G on 24 points as a subgroup of M_{24} splits them into at least five orbits.

Similarly, starting from a subgroup G of M_{24} that acts on 24 points with at least five orbits, one can construct the action of G on $H^{1,1}$. Using the global Torelli theorem, this translates to the existence of a K3 surface S whose symmetry is G.

One example is the Fermat quartic $X^4 + Y^4 + Z^4 + W^4 = 0$ in $\mathbb{CP}^3$. The symmetry is $(\mathbb{Z}_4)^2 \rtimes S_4$, with 384 elements. This is indeed a subgroup of M_{24} that decomposes 24 points into five orbits, of lengths 1, 1, 2, 4, and 16.

More examples and details of the analysis can be found in [Mukai 88, Kondo 98].

ACKNOWLEDGMENTS

The authors thank the hospitality of the Aspen Center of Physics during the workshop “Unity of String Theory,” when the observation reported in this paper was made. We would like to thank S. Mukai for discussions.

T. E. is supported in part by a Grant in Aid from the Japan Ministry of Education, Culture, Sports, Science, and Technology (MEXT). H. O. is supported in part by U.S. Department of Energy grant DE-FG03-92-ER40701, the World Premier International Research Center Initiative, Grant-in-Aid for Scientific Research (C) 20540256 of MEXT, and the Humboldt Research Award. Y. T. is supported in part by NSF grant PHY-0503584, and by a Marvin L. Goldberger membership at the Institute for Advanced Study.

REFERENCES

[Bringmann and Ono 06] K. Bringmann and K. Ono. “The $f(q)$ Mock Theta Function Conjecture and Partition Ranks.” *Invent. Math.* 165 (2006), 243–266.

[Bringmann and Ono 08] K. Bringmann and K. Ono. “Coefficients of Harmonic Maas Forms.” Preprint, 2008.

[Cheng 10] M. C. N. Cheng. “K3 Surfaces, $\mathcal{N}=4$ Dyons, and the Mathieu Group M24.” arXiv:1005.5415 [hep-th], 2010.

[Conway and Norton 79] J. H. Conway and S. P. Norton. “Monstrous Moonshine.” *Bull. London Math. Soc.* 11 (1979), 308–339.

[Conway et al. 85] J. Conway, R. Curtis, S. Norton, R. Parker, and R. Wilson. *Atlas of Finite Groups. Maximal Subgroups and Ordinary Characters for Simple Groups.* Oxford: Clarendon Press, 1985.

[Dong and Mason 94] C. Y. Dong and G. Mason. “An Orbifold Theory of Genus Zero Associated to the Sporadic Group M_{24}.” *Commun. Math. Phys.* 164 (1994) 87–104.

[Eguchi and Hikami 09] T. Eguchi and K. Hikami. “Superconformal Algebras and Mock Theta Functions 2: Rademacher Expansion for K3 Surface.” *Commun. Number Theor. and Phys.* 3 (2009), 531–554.

[Eguchi and Hikami 10] T. Eguchi and K. Hikami, “$N = 4$ Superconformal Algebra and the Entropy of Hyper-Kähler Manifolds.” *JHEP* 1002 (2010), 019.

[Eguchi and Taormina 87] T. Eguchi and A. Taormina. “Unitary Representations of $\mathcal{N}=4$ Superconformal Algebra.” *Phys. Lett. B* 196 (1987), 75–81.

[Eguchi and Taormina 88] T. Eguchi and A. Taormina. “Character Formulas for the $\mathcal{N}=4$ Superconformal Algebra.” *Phys. Lett. B* 200 (1988), 315–322.

[Eguchi et al. 89] T. Eguchi, H. Ooguri, A. Taormina, and S. K. Yang. “Superconformal Algebras and String Compactification on Manifolds with SU(N) Holonomy.” *Nucl. Phys. B* 315 (1989), 193–221.

[Frenkel et al. 88] I. B. Frenkel, J. Lepowsky and A. Meurman. *Vertex Operator Algebras and the Monster*, Pure and Applied Math. 134. New York: Academic Press, 1988.

[Gaberdiel et al. 10] M. R. Gaberdiel, S. Hohenegger, and R. Volpato. “Mathieu Twining Characters for K3.” arXiv:1006.0221 [hep-th], 2010.

[Gannon 04] T. Gannon. “Monstrous Moonshine: The First Twenty-Five Years.” arXiv:math/0402345, 2004.

[Kac 98] V. G. Kac. *Vertex Algebras for Beginners*, 2nd ed., University Lecture Series 10. Providence: American Mathematical Society, 1998.

[Kac and Wakimoto 04] V. G. Kac and M. Wakimoto. “Quantum Reduction and Representation Theory of Superconformal Algebras.” *Advances in Math.* 185 (2004), 400–458.

[Kondo 98]S. Kondo. “Niemeier Lattices, Mathieu Groups and Finite Groups of Symplectic Automorphisms of K3 Surfaces.” *Duke Math. Journal* 92 (1998), 593–603.

[Mukai 88] S. M. “Finite Groups of Automorphisms of K3 Surfaces and the Mathieu Group.” *Invent. Math.* 94 (1988), 183–221.

[Ooguri 89] H. Ooguri. "Superconformal Symmetry and Geometry of Ricci Flat Kähler Manifolds." *Int. J. Mod. Phys. A* 4 (1989), 4303–4324.

[Thompson 79] J. G. Thompson. "Some Numerology between the Fischer–Griess Monster and the Elliptic Modular Function." *Bull. London Math. Soc.* 11 (1979), 352–353.

[Witten 87] E. Witten. "Elliptic Genera and Quantum Field Theory." *Commun. Math. Phys.* 109 (1987), 525–536.

[Witten 07] E. Witten. "Three-Dimensional Gravity Revisited." arXiv:0706.3359, 2007.

[Math. Soc. Japan 07] Mathematical Society of Japan. *Iwanami Suugaku Jiten*, 4th Japanese ed. Tokyo: Iwanami Shoten, 2007. (English translation of the 3rd Japanese edition is available as the 2nd English edition of *Encyclopedic Dictionary of Mathematics*. Cambridge: MIT press, 1987.)

[Zagier 07] D. Zagier, "Ramanujan's Mock Theta Functions and Their Applications." Séminaire Bourbaki 60ème année (2006–2007), no. 986.

[Zwegers 02] S. P. Zwegers. "Mock Theta Functions." Ph.D. thesis, Universiteit Utrecht, 2002.

Tohru Eguchi, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502 Japan (eguchi@yukawa.kyoto-u.ac.jp)

Hirosi Ooguri, California Institute of Technology, Pasadena, CA 91125, USA (ooguri@theory.caltech.edu)

Yuji Tachikawa, School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA (yujitach@ias.edu)

Received April 21, 2010; accepted May 5, 2010.

arXiv:1110.3859v1 [math.RT] 18 Oct 2011

On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine

Miranda C. N. Cheng $^{\flat,\natural}$ and John F. R. Duncan $^{\sharp}$

$^{\flat}$*Department of Mathematics, Harvard University,*
Cambridge, MA 02138, U.S.A.
$^{\natural}$*Department of Physics, Harvard University,*
Cambridge, MA 02138, U.S.A.
$^{\sharp}$ *Department of Mathematics, Case Western Reserve University,*
Cleveland, OH 44106, U.S.A.

Abstract

Recently a conjecture has been proposed which attaches (mock) modular forms to the largest Mathieu group. This may be compared to monstrous moonshine, in which modular functions are attached to elements of the Monster group. One of the most remarkable aspects of monstrous moonshine is the following genus zero property: the modular functions turn out to be the generators for the function fields of their invariance groups. In particular, these invariance groups define genus zero quotients of the upper half plane. It is therefore natural to ask if there is an analogue of this property in the Mathieu case, and at first glance the answer appears to be negative since not all the discrete groups arising there have genus zero. On the other hand, in this article we prove that each (mock) modular form appearing in the Mathieu correspondence coincides with the Rademacher sum constructed from its polar part. This property, inspired by the AdS/CFT correspondence in physics, was shown previously to be equivalent to the genus zero property of monstrous moonshine. Hence we conclude that this "Rademacher summability" property serves as the natural analogue of the genus zero property in the Mathieu case. Our result constitutes further evidence that the Rademacher method provides a powerful framework for understanding the modularity of moonshine, and leads to interesting physical questions regarding the gravitational duals of the relevant conformal field theories.

Contents

1 Introduction 2

2 The Largest Mathieu Group 5

3 Mock Modular Forms 7

4 Rademacher Sums 10

5 Convergence 13

6 Coefficients 15

7 Variance 17

8 Coincidence 20

9 Spectral Theory 29

10 Conclusion and Discussion 32

A Dedekind Eta Function 35

B Bessel Function 35

C Lipschitz Summation 36

D Some Details on Convergence 37

E Rademacher Sums at Other Cusps 38

F Modular Forms and Representations 40

1 Introduction

In 2010, an unexpected relation between the elliptic genus of a $K3$ surface and the sporadic group M_{24} was observed by Eguchi, Ooguri and Tachikawa [1]. Subsequently, the connection between the $K3$ elliptic genus twisted by a group element of M_{24} and the characters of M_{24}-representations was first explored in [2, 3], and further studied in [4, 5]. From the point of view of the $K3$ superconformal sigma model, these twisted (or twining) elliptic genera are believed to specify the way in which the $K3$ elliptic cohomology furnishes an M_{24}-representation of infinite dimension. Notice that such a twisted object depends only on the conjugacy class $[g]$ to which the element g belongs. From the $\mathcal{N} = 4$ superconformal symmetry of the sigma model it follows that they are weak Jacobi forms, and moreover admit natural decompositions leading to a set of (mock) modular forms H_g of weight 1/2. As a result of these studies the existence of a natural relationship between representations of M_{24} and the mock modular forms H_g has been conjectured. While such a relationship is yet to be established, extensive evidence has been gathered in [2, 4, 5]. In this context the H_g are often referred to as the McKay–Thompson series and a key prediction of the conjecture is that they exhibit particular modular behaviour under certain discrete subgroups Γ_g of $\mathrm{SL}_2(\mathbb{R})$.

This may be compared to the moonshine conjecture of Conway and Norton [6], which relates genus zero subgroups of $\mathrm{SL}_2(\mathbb{R})$ to elements of the Monster group. More precisely, to each conjugacy class $[g]$ of the Monster is attached a function T_g (its McKay–Thompson series) whose Fourier coefficients are given by characters of the Monster. The genus zero property of monstrous moonshine then refers to the empirical fact (ultimately proven by Borcherds [7]) that T_g defines a function on the upper half plane that generates the function field of its invariance group Γ_g. In particular, the (compactified) quotient of the upper half plane by Γ_g is a Riemann surface of genus zero. This property is also expected to extend to the so-called generalised moonshine [8] which attaches modular functions to commuting pairs of elements of the Monster. Ogg's observation [9] that the normaliser of $\Gamma_0(p)$ in $\mathrm{SL}_2(\mathbb{R})$ for p a prime is genus zero if and only if p divides the order of the Monster was arguably the first hint of a relation between the Monster and genus zero groups.

Assuming the validity of the M_{24} conjecture, and in particular, the validity of the descriptions of the H_g given in [2, 4, 5], a natural question to ask is whether the groups $\Gamma_g \subset SL_2(\mathbb{R})$ attached to the McKay–Thompson series H_g for M_{24} share this fundamental pattern present in monstrous moonshine: do they have a similar genus zero property? At first glance the answer is negative since not all the groups Γ_g arising from M_{24} define genus zero quotients of the upper half plane. When a pattern breaks, we may seek to replace it with a new pattern which encompasses both the old paradigm and the new "exception". Moreover, we also want to understand why a pattern exists at all in the first place. These are exactly the goals of this paper, and these considerations lead us to recast the question in a new form: a form suggested by earlier work of I. Frenkel which reconsidered the origin, and interpretation, of the genus zero property of monstrous moonshine.

The article [10] makes an important step towards a physical explanation of the genus zero property. There it is shown that a discrete group Γ has genus zero if and only if a certain regularisation procedure, inspired by Rademacher's work on the j-function [11] and summarised here as "Reg", renders the Poincaré-like series

$$T_\Gamma(\tau) = \mathrm{Reg}\Big(\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} q^{-1}\big|_\gamma \Big) \tag{1.1}$$

convergent without spoiling invariance under the group Γ. (We say a group Γ has genus zero if it defines a genus zero quotient of the upper half-plane.) In the above formula Γ_∞ is the subgroup of Γ

fixing the cusp (representative) $i\infty$, the sum is over a set of representatives γ for the cosets of Γ_∞ in Γ, and $q = e^{2\pi i\tau}$. As is turns out, the above expression, which we refer to as a *Rademacher sum*, and which can be used to define a function for any discrete subgroup Γ of $\mathrm{SL}_2(\mathbb{R})$ commensurable with $\mathrm{SL}_2(\mathbb{Z})$ (for example), is anomaly free—that is, invariant under Γ even after the regularisation—if and only if the group Γ has genus zero. In the case that Γ does have genus zero the Rademacher sum (1.1) defines a generator for the field of Γ-invariant functions, and it defines precisely the generator T_g when $\Gamma = \Gamma_g$ for g in the Monster. So the genus zero property of monstrous moonshine may be reformulated as follows:

> The McKay–Thompson series T_g coincides with the Rademacher sum T_Γ attached to its invariance group $\Gamma = \Gamma_g$.

Significantly, Rademacher sums with a higher order pole q^{-n}, which are relevant for conformal field theories (CFTs) with higher central charges, also define modular functions as long as the group Γ has genus zero.

A compelling physical interpretation of the Rademacher sum is provided by the so-called AdS/CFT correspondence [12] (also referred to as the gauge/gravity duality or the holographic duality in more general contexts), which asserts, among many other things, that the partition function of a given two dimensional CFT "with an AdS dual" equals the partition function of another physical theory in three Euclidean dimensions with gravitational interaction and with asymptotically anti de Sitter (AdS) boundary condition and toroidal conformal boundary. There is not yet a systematic understanding of the conditions a CFT has to satisfy in order for it to have an AdS dual. Nevertheless, the correspondence, when applicable, provides both deep intuitive insights and powerful computational tools for the study of the theory. From the fact that the only smooth three-manifold with asymptotically AdS torus boundary condition is a solid torus, it follows that the saddle points of such a partition function are labeled by the different possible ways to "fill in the torus"; that is, the different choices of primitive cycle on the boundary torus which may become contractible in a solid torus that fills it [13]. These different saddle points are therefore labeled by the coset space $\Gamma_\infty\backslash\Gamma$, where $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ [14]. From a bulk, gravitational point of view, the group $\mathrm{SL}_2(\mathbb{Z})$ has an interpretation as the group of large diffeomorphisms, and Γ_∞ is the subgroup that leaves the contractible cycle invariant and therefore can be described by a mere change of coordinates. Such considerations underlie the previous use of Rademacher sums in the physics literature [14, 15, 16, 17, 18, 19].

Apart from the partition function, which computes the dimension of the Hilbert space graded by the basic quantum numbers (the energy, for instance) of the theory, more interesting information can be gained in the presence of a non-trivial symmetry group acting on the Hilbert space by studying the twisted partition function (a trace over the Hilbert space with a group element inserted) which computes the group characters of the Hilbert space. In the Lagrangian formulation of quantum field theories this twisting corresponds to a modification of the boundary condition. Now, in the study of moonshine the McKay–Thompson series attached to the group element g has a physical interpretation as the twisted partition function of a two dimensional CFT with a boundary condition twisted by g. This twisting procedure has a clear geometrical meaning for a two dimensional CFT with an AdS gravity dual: From the point of view of the gravity theory, the change of boundary condition by insertion of a group element g also changes the set of allowed saddle points, and relatedly, the allowed large diffeomorphisms is now given by a different discrete group $\Gamma_g \subset \mathrm{SL}_2(\mathbb{R})$. From this argument, assuming that the moonshine CFT (or its cousins with higher central charges) have semi-classical-like AdS descriptions in which the path integrals are controlled by their saddle points, the modular objects should admit expressions as Rademacher sums over the coset spaces $\Gamma_\infty\backslash\Gamma_g$. This picture motivates the relevance of the Rademacher summability property:

> We say a function f is *Rademacher summable* if it coincides with the Rademacher sum attached to its polar part.

As we have observed above, Rademacher summability is equivalent to the genus zero property in the monstrous case by the results of [10], where the McKay-Thompson series are modular functions. Moreover, the Rademacher summing procedure applied to Γ produces functions which do not have singularities at cusps of Γ other than the infinite one. This is in accordance with our expectation that our theory of gravity should admit a unique low energy description.

The situation is arguably more subtle and also more interesting for M_{24}. As mentioned earlier, in this case the relevant modular objects are weight 1/2 (mock) modular forms, and one does not expect any clear relation between the genus of Γ_g and the Rademacher summability of H_g—indeed, not all the Γ_g attached to elements of M_{24} have genus zero—while in this work we verify explicitly that the functions H_g are all Rademacher summable:

$$H_g(\tau) = -2\,\mathrm{Reg}\Big(\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_g} q^{-1/8}\big|_\gamma\Big) \tag{1.2}$$

for all $g \in M_{24}$. The precise meaning of the above equation will be elucidated in §4, and is in particular the content of the main theorem (Theorem 4.1) of the present paper. Closely related sums for two of the above (mock) modular forms, corresponding to taking g to be the identity or in a certain class of order 2, have been suggested previously in [20]. The function H_g for g the identity is one of the mock modular forms constructed over a decade ago by Pribitkin in [21].

Note the remarkable similarity between (1.2) and (1.1), and recall that Rademacher summability is equivalent to the genus zero property in the monstrous case. In view of this, the Rademacher summability property clearly serves as a natural analogue of the genus zero property, and as a new paradigm that applies to both the Monster and M_{24}. Moreover, Rademacher sums provide a new angle from which to view the moonshine phenomenon: is it not miraculous that the Rademacher machinery, applied with such simple inputs, uniformly produces infinite q-series that contain so much information about sporadic simple groups?

The rest of the paper is organised as follows. In §2 we review the basic properties of the sporadic group M_{24}, focusing on its 24-dimensional permutation representation. In §3 we review the properties of the (mock) modular forms that have been conjectured to encode an infinite-dimensional, $\mathbb{Z}$-graded M_{24} module. In §4 we explain our construction of the Rademacher sums and state the main theorem (Theorem 4.1) of the present paper. The rest of the paper is devoted to the proof of Theorem 4.1. In §5 we discuss the convergence of the Rademacher sums and establish that these sums are indeed convergent given the convergence of certain Selberg–Kloosterman zeta functions, which will be shown in §9. Having established the convergence of the sums, in §6 we give explicit expressions for the Fourier coefficients of our Rademacher sums, and discuss their asymptotic growth. In §7 we study the transformation properties of the Rademacher sums and establish their (mock) modularity. Given the modularity and information about their behaviour at the cusps, in §8 we finally establish that the Rademacher sums constructed in §4 indeed coincide with the (mock) modular forms which are the proposed McKay–Thompson series for the group M_{24} described in §3.

2 The Largest Mathieu Group

We shall start by recalling some facts about the largest Mathieu group. The group M_{24} may be characterised as the automorphism group of the unique doubly even self-dual binary code of length 24 with no words of weight 4, also known as the *(extended) binary Golay code*. In other words, there is a unique (up to permutation) set, $\mathcal{G}$ say, of length 24 binary codewords (sequences of 0's and 1's) such that any other length 24 codeword has even overlap with all the codewords of $\mathcal{G}$ if and only if this word itself is in $\mathcal{G}$, and the number of 1's in each codeword of $\mathcal{G}$ is divisible by and not equal to 4. The group of permutations of the 24 coordinates that preserves the set $\mathcal{G}$ is the sporadic group M_{24}. See, for instance, [22].

$[g]$	cycle shape	$\eta_g(\tau)$	k_g	n_g	N_g	h_g
$1A$	1^{24}	$\eta^{24}(\tau)$	12	1	1	1
$2A$	$1^8 2^8$	$\eta^8(\tau)\eta^8(2\tau)$	8	2	2	1
$2B$	2^{12}	$\eta^{12}(2\tau)$	6	2	4	2
$3A$	$1^6 3^6$	$\eta^6(\tau)\eta^6(3\tau)$	6	3	3	1
$3B$	3^8	$\eta^8(3\tau)$	4	3	9	3
$4A$	$2^4 4^4$	$\eta^4(2\tau)\eta^4(4\tau)$	4	4	8	2
$4B$	$1^4 2^4 4^4$	$\eta^4(\tau)\eta^2(2\tau)\eta^4(4\tau)$	5	4	4	1
$4C$	4^6	$\eta^6(4\tau)$	3	4	16	4
$5A$	$1^4 5^4$	$\eta^4(\tau)\eta^4(5\tau)$	4	5	5	1
$6A$	$1^2 2^2 3^2 6^2$	$\eta^2(\tau)\eta^2(2\tau)\eta^2(3\tau)\eta^2(6\tau)$	4	6	6	1
$6B$	6^4	$\eta^4(6\tau)$	2	6	36	6
$7AB$	$1^3 7^3$	$\eta^3(\tau)\eta^3(7\tau)$	3	7	7	1
$8A$	$1^2 2^1 4^1 8^2$	$\eta^2(\tau)\eta(2\tau)\eta(4\tau)\eta^2(8\tau)$	3	8	8	1
$10A$	$2^2 10^2$	$\eta^2(2\tau)\eta^2(10\tau)$	2	10	20	2
$11A$	$1^2 11^2$	$\eta^2(\tau)\eta^2(11\tau)$	2	11	11	1
$12A$	$2^1 4^1 6^1 12^1$	$\eta(2\tau)\eta(4\tau)\eta(6\tau)\eta(12\tau)$	2	12	24	2
$12B$	12^2	$\eta^2(12\tau)$	1	12	144	12
$14AB$	$1^1 2^1 7^1 14^1$	$\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau)$	2	14	14	1
$15AB$	$1^1 3^1 5^1 15^1$	$\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau)$	2	15	15	1
$21AB$	$3^1 21^1$	$\eta(3\tau)\eta(21\tau)$	1	21	63	3
$23AB$	$1^1 23^1$	$\eta(\tau)\eta(23\tau)$	1	23	23	1

Table 1: The cycle shapes, weights (k_g), levels (N_g) and orders (n_g) of the 26 conjugacy classes of the sporadic group M_{24}. The length of the shortest cycle is $h_g = N_g/n_g$. The naming of the conjugacy classes follows the ATLAS convention (cf. [23]). We write $7AB$, for example, to indicate that the entries of the incident row are valid for both the conjugacy classes $7A$ and $7B$.

As such, M_{24} naturally admits a permutation representation of degree 24, and this allows us to assign a *cycle shape* to each of its elements. For example, to the identity element we associate the cycle shape 1^{24}; to an element of M_{24} that is a product of 12 mutually commuting transpositions we associate the cycle shape 2^{12}, and so on. More generally, any cycle shape arising from an element of M_{24} (or S_{24}, for that matter) is of the form

$$i_1{}^{\ell_1} i_2{}^{\ell_2} \cdots i_r{}^{\ell_r}, \quad \sum_{s=1}^{r} \ell_s \, i_s = 24 \, ,$$

for some $\ell_s \in \mathbb{Z}^+$ and $1 \leq i_1 < \cdots < i_r \leq 23$ with $r \geq 1$. Clearly, the cycle shape of an element of M_{24} depends only on its conjugacy class, denoted by $[g]$, although different conjugacy classes can share the same cycle shape. For future reference we denote the character underlying this defining 24-dimensional representation of M_{24} by χ. The value of the character $\chi(g)$ equals the number of fixed points of g in the action on the set of 24 points. In particular, note that $\chi(g) = \ell_1$ in case $i_1 = 1$ and $\chi(g) = 0$ otherwise.

It turns out that the cycle shapes of M_{24} have many special properties that will be important for the understanding of the modular properties of the associated McKay–Thompson series which we will discuss shortly. First, the M_{24} cycle shapes are privileged in that they are all of the so-called *balanced type* (cf. [6]), meaning that for each $g \in M_{24}$ there exists a positive integer N_g such that if $\prod i_s^{\ell_s}$ is the cycle shape associated to g then

$$\prod_s i_s^{\ell_s} = \prod_s \left(\frac{N_g}{i_s}\right)^{\ell_s} .$$

We will refer to the number N_g as the *level* of the g.

If g has cycle shape $i_1^{\ell_1} \cdots i_r^{\ell_r}$ then the order of g is the least common multiple of the i_s's. A second special property of $[g] \subset M_{24}$ is that the order of g coincides with the length i_r of the longest cycle in the cycle shape. Henceforth we will denote $n_g = i_r$.

Finally, observe that for all $g \in M_{24}$ the level N_g defined above equals the product of the shortest and the longest cycle. Hence we have $h_g n_g = N_g$ where h_g denotes the length of the shortest cycle in the cycle shape. Moreover, we also have the property $h_g | n_g$ and $h_g | 12$. This is very reminiscent of the monstrous moonshine [6]. We also set $k_g = \sum_{s=1}^{r} \ell_s / 2$ to be half of the total number of cycles and call it the *weight* of g. Of course, N_g, n_g, h_g and k_g depend only on the conjugacy class $[g]$ containing g and can be found in Table 1.

To each element $g \in M_{24}$ we can attach an *eta-product*, to be denoted η_g, which is the function on the upper half-plane given by

$$\eta_g(\tau) = \prod_s \eta(i_s \tau)^{\ell_s} \tag{2.1}$$

where $\prod_s i_s^{\ell_s}$ is the cycle shape attached to g, and $\eta(\tau)$ is the Dedekind eta function satisfying $\eta(\tau) = q^{1/24} \prod_{n \in \mathbb{Z}^+} (1 - q^n)$ for $q = e(\tau)$, where for later convenience, here and everywhere else in this paper we will use the shorthand notation

$$e(x) = e^{2\pi i x} .$$

As was observed in [24, 25], the eta-product η_g associated to an element $g \in M_{24}$ (or rather, to its conjugacy class $[g]$) is a cusp form of weight k_g for the group $\Gamma_0(N_g)$, with a *Dirichlet character* ς_g that is trivial if the weight k_g is even and is otherwise defined, in terms of the Jacobi symbol $(\frac{n}{m})$, by

$$\varsigma_g(\gamma) = \begin{cases} \left(\frac{N_g}{d}\right)(-1)^{\frac{d-1}{2}}, & d \text{ odd,} \\ \left(\frac{N_g}{d}\right), & d \text{ even,} \end{cases}$$

in case d is the lower right entry of $\gamma \in \Gamma_0(N_g)$. Let's recall that

$$\Gamma_0(N) = \left\{ \gamma \middle| \ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,, \ c = 0 \mod N \right\} .$$

The eta-product η_g also defines a cusp form of weight k_g on the larger (or equal) group $\Gamma_0(n_g) \supseteq \Gamma_0(N_g)$ if we allow for a slightly more sophisticated multiplier system. We remark here that, according to our conventions, a function $\xi : \Gamma \to \mathbb{C}^*$ is called a *multiplier system for* Γ *of weight* w in case the identity

$$\xi(\gamma\sigma)\,\mathrm{jac}(\gamma\sigma,\tau)^{w/2} = \xi(\gamma)\xi(\sigma)\,\mathrm{jac}(\gamma,\sigma\tau)^{w/2}\,\mathrm{jac}(\sigma,\tau)^{w/2} \tag{2.2}$$

holds for all $\gamma,\sigma \in \Gamma$ and $\tau \in \mathbb{H}$. Writing $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have in the above formula $\gamma\tau = \frac{a\tau+b}{c\tau+d}$, and the Jacobian

$$\mathrm{jac}(\gamma,\tau) = (c\tau+d)^{-2}\ .$$

In detail, for $\gamma \in \Gamma$, we define the *slash operator* $|_{\xi,w}$ *of weight* w *associated with multiplier* ξ as

$$(f|_{\xi,w}\gamma)(\tau) = \xi(\gamma)f(\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{w/2}\ , \tag{2.3}$$

and in particular a holomorphic function $f : \mathbb{H} \to \mathbb{C}$ is a modular form of weight w and multiplier ξ on Γ if and only if $(f|_{\xi,w}\gamma)(\tau) = f(\tau)$ for all $\gamma \in \Gamma$. In the present case we have

$$\Big(\frac{1}{\eta_g}\Big|_{\xi_g,k_g}\gamma\Big)(\tau) = \frac{1}{\eta_g(\tau)} \quad , \quad \text{for all} \quad \gamma \in \Gamma_0(n_g)\ , \tag{2.4}$$

where $\xi_g(\gamma) = \rho_{n_g|h_g}(\gamma)\varsigma_g(\gamma)$ with

$$\rho_{n|h}(\gamma) = \mathrm{e}(-\tfrac{1}{(\gamma\infty-\gamma 0)nh}) = \mathrm{e}(-cd/nh)\ . \tag{2.5}$$

Note that $\rho_{n_g|h_g}$ is actually a character on $\Gamma_0(n_g)$ since we have that $xy \equiv 1 \pmod{h_g}$ implies $x \equiv y \pmod{h_g}$ by virtue of the fact that $h_g = N_g/n_g$ is a divisor of 24 for every $g \in M_{24}$ (cf. [6, §3]). Evidently the kernel of $\rho_{n|h}$ is $\Gamma_0(nh)$, and in particular, $\rho_{n|h}$ is trivial on Γ_∞.

These (meromorphic) modular forms $1/\eta_g(\tau)$ have also the interpretation as the partition function of a the conformal field theory of 24 free chiral bosons, twisted by $g \in M_{24}$ which acts on the 24 bosons as an element of S_{24}. As such, they are also McKay–Thompson series whose Fourier coefficients are positive-integral linear combinations of M_{24} characters. See Table 4 for the first ten M_{24}-representations appearing in $1/\eta_g(\tau)$. Following an earlier observation in [26], it was shown in [2] that they are connected to the elliptic genera of $K3$ surfaces and the moonshine for M_{24} which will be discussed in the next section, via a lifting to Siegel modular forms and generalised Kac-Moody superalgebras. More details will be given in [27].

Observe that the behaviour of the multipliers naturally divides the conjugacy classes of M_{24} into two types: those for which the phase is trivial on $\Gamma_0(n_g)$ and those for which it is not. The former occurs just when $h_g = 1$, and the elements of $g \in M_{24}$ with $h_g > 1$ are exactly the classes which act fixed-point-freely in the defining permutation representation on 24 points. The elements with $h_g = 1$ have at least one fixed point and thus can be located in a (maximal) subgroup of M_{24} isomorphic to the second largest Mathieu group, M_{23}.

3 Mock Modular Forms

In the recent article [1] a remarkable observation relating M_{24} and the unique (up to scale) weak Jacobi form of weight zero and index one, here denoted $Z(\tau, z)$, was made via a decomposition of the latter object into a combination of mock theta series and mock modular forms. As is shown in [28, 29, 20], the function $Z(\tau,z)=8\sum_{i=2,3,4}(\frac{\theta_i(\tau,z)}{\theta_i(\tau,0)})^2$ admits an expression

$$Z(\tau,z)=\frac{\theta_1(\tau,z)^2}{\eta(\tau)^3}\left(a\,\mu(\tau,z)+q^{-1/8}\big(b+\sum_{k=1}^{\infty}t_k\,q^k\big)\right) \tag{3.1}$$

with $a,b,t_k\in\mathbb{Z}$ for all $k\in\mathbb{Z}^+$ where $\mu(\tau,z)$ denotes the *Appell-Lerch sum*, satisfying

$$\mu(\tau,z)=\frac{-iy^{1/2}}{\theta_1(\tau,z)}\sum_{\ell=-\infty}^{\infty}\frac{(-1)^\ell y^n q^{\ell(\ell+1)/2}}{1-yq^\ell}$$

for $q=e(\tau)$ and $y=e(z)$. The Jacobi theta functions are given by

$$\begin{aligned}\theta_1(\tau,z)&=-iq^{1/8}y^{1/2}\prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^{n-1})\\ \theta_2(\tau,z)&=q^{1/8}y^{1/2}\prod_{n=1}^{\infty}(1-q^n)(1+yq^n)(1+y^{-1}q^{n-1})\\ \theta_3(\tau,z)&=\prod_{n=1}^{\infty}(1-q^n)(1+y\,q^{n-1/2})(1+y^{-1}q^{n-1/2})\\ \theta_4(\tau,z)&=\prod_{n=1}^{\infty}(1-q^n)(1-y\,q^{n-1/2})(1-y^{-1}q^{n-1/2})\ .\end{aligned} \tag{3.2}$$

By inspection, $a=24=\chi(1A)$, $b=-2$ and the first few t_k are

$$2\times 45,\ 2\times 231,\ 2\times 770,\ 2\times 2277,\ 2\times 5796\cdots$$

These positive integers t_k have the interpretation of enumerating the k-th massive representations of the $\mathcal{N}=4$ superconformal algebra in the elliptic genus of a $K3$ surface.

The surprising connection to M_{24}, beyond the fact that $a=24$ is the dimension of the defining permutation representation of M_{24}, is the following: the integers 45, 231, 770, 2277 and 5796, which are the $t_k/2$ for $k=1,2,3,4,5$, are the dimensions of irreducible representations of M_{24}. It was conjectured that the entire set of values t_k for $k\in\mathbb{Z}^+$ encode the graded dimension of a naturally defined graded M_{24} module $K=\bigoplus_k K_k$ with $\dim K_k=t_k$. If the conjecture holds then we can obtain new functions by replacing t_k with $\mathrm{tr}_{K_k}g$ for $g\in M_{24}$. This idea was first investigated in [2], and independently in [3]; see also [4] and [5]. It is an idea that we pursue further in this article.

Define $H(\tau)$ so that $Z(\tau,z)\eta(\tau)^3=\theta_1(\tau,z)^2(a\mu(\tau,z)+H(\tau))$. Then

$$H(\tau)=q^{-\frac{1}{8}}\left(-2+\sum_{k=1}^{\infty}t_kq^k\right) \tag{3.3}$$

so that $H(\tau)$ is (essentially) a power series incorporating the t_k, and hence the graded dimension of the conjectural M_{24} module K. This function $H(\tau)$ enjoys a special relationship with the group $\mathrm{SL}_2(\mathbb{Z})$; namely, it is a *weakly holomorphic mock modular form of weight* $1/2$ on $\mathrm{SL}_2(\mathbb{Z})$ with *shadow* $24\,\eta(\tau)^3$ (cf. [30]), which means that $H(\tau)$ is a holomorphic function on the upper half-plane $\mathbb{H}$

with at most exponential growth as $\tau \to \alpha$ for any $\alpha \in \mathbb{Q}$, and if we define the *completion* of $H(\tau)$, to be denoted $\hat{H}(\tau)$, by setting

$$\hat{H}(\tau) = H(\tau) + 24\,(4i)^{-1/2} \int_{-\bar{\tau}}^{\infty} (z+\tau)^{-1/2}\overline{\eta(-\bar{z})^3}\mathrm{d}z,$$

then $\hat{H}(\tau)$ transforms as a modular form of weight $1/2$ on $\mathrm{SL}_2(\mathbb{Z})$ with multiplier system conjugate to that of $\eta(\tau)^3$. In other words, we have

$$\big(\hat{H}(\tau)|_{\epsilon^{-3},1/2}\gamma\big)(\tau) = \epsilon(\gamma)^{-3}\hat{H}(\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{1/4} = \hat{H}(\tau)$$

for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, where $\epsilon : \mathrm{SL}_2(\mathbb{Z}) \to \mathbb{C}^*$ is the multiplier system for $\eta(\tau)$ satisfying

$$\big(\eta|_{\epsilon,1/2}\gamma\big)(\tau) = \eta(\tau)\ .$$

(See §A for an explicit description of ϵ.)

More generally, a holomorphic function $h(\tau)$ on $\mathbb{H}$ is called a *(weakly holomorphic) mock modular form of weight w* for a discrete group Γ (e.g. a congruence subgroup of $\mathrm{SL}_2(\mathbb{R})$) if it has at most exponential growth as $\tau \to \alpha$ for any $\alpha \in \mathbb{Q}$, and if there exists a holomorphic modular form $g(\tau)$ of weight $2-w$ on Γ such that $\hat{h}(\tau)$, given by

$$\hat{h}(\tau) = h(\tau) + (4i)^{w-1} \int_{-\bar{\tau}}^{\infty} (z+\tau)^{-w}\overline{g(-\bar{z})}\mathrm{d}z, \tag{3.4}$$

is a (non-holomorphic) modular form of weight w for Γ for some multiplier system ψ say. In this case the function g is called the *shadow* of the mock modular form h and we call ψ the multiplier system of h.

If the conjectural M_{24}-module K exists, apart from $H(\tau)$ there must be a family of functions $H_g(\tau)$, for each conjugacy class $[g] \subset M_{24}$, obtainable by replacing each t_k with the trace of g on K_k. These are the McKay–Thompson series given by

$$H_g(\tau) = q^{-\frac{1}{8}}\left(-2 + \sum_{k=1}^{\infty} \mathrm{tr}_{K_k}(g) q^k\right)\ . \tag{3.5}$$

Explicit expressions for these function $H_g(\tau)$ have been proposed in [2, 3], mostly for conjugacy classes with $h_g = 1$, and completed for all $[g] \subset M_{24}$ in [4, 5]. These proposals state that they are given by $H(\tau)$, the character $\chi(g)$ of g in the 24-dimensional representation, and certain weight two modular forms $\tilde{T}_g(\tau)$ for the group $\Gamma_0(N_g)$, by

$$H_g(\tau) = \frac{\chi(g)}{24}H(\tau) - \frac{\tilde{T}_g(\tau)}{\eta(\tau)^3}\ . \tag{3.6}$$

These data are collected in Table 2. The expression (3.6) makes manifest that the q-series $H_g(\tau)$ is a mock modular form with shadow $\chi(g)\eta^3(\tau)$, and is in particular a usual modular form when $h_g > 1$, or equivalently, $\chi(g) = 0$. Moreover, it is easy to check that $H_g(\tau)$ transforms nicely under the group $\Gamma_0(n_g)$, with the multiplier system $\psi(\gamma) = \epsilon(\gamma)^{-3}\rho_{n_g|h_g}(\gamma)$. Notice that the extra multiplier ρ that appears when $h_g \neq 1$ is the same as that of the inverse eta-products $1/\eta_g(\tau)$ which are also related to M_{24}, a fact that is in accordance with the 1/2- and 1/4-BPS spectrum of the

$[g]$	$\chi(g)$	$\tilde{T}_g(\tau)$
$1A$	24	0
$2A$	8	$16\Lambda_2$
$2B$	0	$-24\Lambda_2+8\Lambda_4=2\eta(\tau)^8/\eta(2\tau)^4$
$3A$	6	$6\Lambda_3$
$3B$	0	$2\eta(\tau)^6/\eta(3\tau)^2$
$4A$	0	$4\Lambda_2-6\Lambda_4+2\Lambda_8=2\eta(2\tau)^8/\eta(4\tau)^4$
$4B$	4	$4(-\Lambda_2+\Lambda_4)$
$4C$	0	$2\eta(\tau)^4\eta(2\tau)^2/\eta(4\tau)^2$
$5A$	4	$2\Lambda_5$
$6A$	2	$2(-\Lambda_2-\Lambda_3+\Lambda_6)$
$6B$	0	$2\eta(\tau)^2\eta(2\tau)^2\eta(3\tau)^2/\eta(6\tau)^2$
$7AB$	3	Λ_7
$8A$	2	$-\Lambda_4+\Lambda_8$
$10A$	0	$2\eta(\tau)^3\eta(2\tau)\eta(5\tau)/\eta(10\tau)$
$11A$	2	$2(\Lambda_{11}-11\eta(\tau)^2\eta(11\tau)^2)/5$
$12A$	0	$2\eta(\tau)^3\eta(4\tau)^2\eta(6\tau)^3/\eta(2\tau)\eta(3\tau)\eta(12\tau)^2$
$12B$	0	$2\eta(\tau)^4\eta(4\tau)\eta(6\tau)/\eta(2\tau)\eta(12\tau)$
$14AB$	1	$(-\Lambda_2-\Lambda_7+\Lambda_{14}-14\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau))/3$
$15AB$	1	$(-\Lambda_3-\Lambda_5+\Lambda_{15}-15\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau))/4$
$21AB$	0	$(7\eta(\tau)^3\eta(7\tau)^3/\eta(3\tau)\eta(21\tau)-\eta(\tau)^6/\eta(3\tau)^2)/3$
$23AB$	1	$(\Lambda_{23}-23\phi_{23,1}+23\phi_{23,2})/11$

Table 2: In this table we collect the data that via equation (3.6) define the weight 1/2 (mock) modular forms $H_g(\tau)$. For $N \in \mathbb{Z}_+$, we denote by Λ_N the weight 2 modular form on $\Gamma_0(N)$ given by $\Lambda_N = Nq\frac{d}{dq}(\log\frac{\eta(N\tau)}{\eta(\tau)})$ (cf. (8.13)). For $N=23$, there are two new forms and we choose the basis $\phi_{23,1}=\eta^2_{23AB}$ and $\phi_{23,2}$ given in (8.14).

$\mathcal{N}=4$, d=4 theory obtained by $K3\times T^2$ compactification of the type II string theory. These mock modular forms $H_g(\tau)$ are the central objects of study in this work.

Although the conjecture (3.5) stating that the Fourier coefficients of $H_g(\tau)$ are all given by characters of M_{24}-module still remains to be proven, it has been checked up to the first 1000 terms. We list the first few in Table 5. Together with other pieces of evidence, this makes it an extremely plausible conjecture. In this paper we will hence assume the validity of this conjecture as a part of our motivation, while our main result (Theorem 4.1) holds independent of it.

4 Rademacher Sums

In this section we will state and explain the main theorem (Theorem 4.1) of the paper, while postponing the proof till the later sections.

Recall that $H(\tau)$ is a mock modular form of weight 1/2 on $\mathrm{SL}_2(\mathbb{Z})$ with a multiplier system given by $\epsilon(\gamma)^{-3}$ and with leading term $-2q^{-1/8}$ in its q-expansion. We may consider the problem of attaching a function with these data via Rademacher sums, and also analogous functions to discrete subgroups of $\mathrm{SL}_2(\mathbb{R})$ other than the modular group[1].

---

[1] For the purpose of establishing relations to M_{24} it suffices to look at subgroups of $\mathrm{SL}_2(\mathbb{Z})$, while it is possible to

Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ containing the group Γ_∞ of upper triangular matrices in $\mathrm{SL}_2(\mathbb{Z})$. In particular, we assume that Γ contains $-\mathrm{Id}$. Define $R_\Gamma(\tau)$ by setting

$$R_\Gamma(\tau) = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \epsilon(\gamma)^{-3}\,\mathrm{e}(-\tfrac{\gamma\tau}{8})\,\mathrm{reg}(\gamma,\tau)\,\mathrm{jac}(\gamma,\tau)^{1/4}\ , \tag{4.1}$$

where $\mathrm{reg}(\gamma,\tau)$ is the *Rademacher regularisation factor (of weight* $1/2$ *and index* $-1/8$*)*, defined by $\mathrm{reg}(\gamma,\tau)=1$ in case γ is upper triangular (i.e. $\gamma\cdot\infty=\infty$) and

$$\mathrm{reg}(\gamma,\tau) = \mathrm{e}(\tfrac{\gamma\tau-\gamma\infty}{8})\ \mathrm{e}(\tfrac{\gamma\infty-\gamma\tau}{8},\tfrac{1}{2}) \tag{4.2}$$

otherwise, where $\mathrm{e}(x,s)$ is the following generalisation of the exponential function:

$$\mathrm{e}(x,s) = \sum_{m\geq 0}\frac{(2\pi i x)^{m+s}}{\Gamma(m+s+1)}\ . \tag{4.3}$$

Note that we recover $\mathrm{e}(x)=e^{2\pi i x}$ by taking $s=0$ in $\mathrm{e}(x,s)$, and for n a positive integer $\mathrm{e}(x,n)$ is the difference between $\mathrm{e}(x)$ and the order $n-1$ Taylor approximation to $\mathrm{e}(x)$. It is also closely related to the incomplete Gamma function $\Gamma(s,x)=\int_x^\infty t^{s-1}e^t\mathrm{d}t$: applying integration by parts repeatedly, we obtain

$$\mathrm{e}(x,s) = \frac{\mathrm{e}(x)}{\Gamma(s)}\int_0^{2\pi i X} t^{s-1}e^t\mathrm{d}t\ . \tag{4.4}$$

Let T be the element of Γ_∞ such that $\mathrm{tr}(T)=2$ and $T\tau=\tau+1$. Note that two elements γ,σ of Γ are in the same right coset of Γ_∞ if and only if there is some $n\in\mathbb{Z}$ such that $\sigma=T^n\gamma$ or $-T^n\gamma$. In particular, they have the same lower rows up to multiplication by $\pm\mathrm{Id}$ and we can therefore use the lower rows to parametrise the cosets. This motivates the sum over the "rectangle" in (4.1): For any given positive K, we consider a sum over a set $\{\gamma\}$ of representatives for the cosets of Γ_∞ in Γ whose elements have lower rows (c,d) satisfying $0\leq c<K$ and $-K^2<d<K^2$. Using a superscript $^\times$ to indicate a sum over coset representatives other than the trivial one, we have

$$(\Gamma_\infty\backslash\Gamma)^\times_{<K} = \left\{\Gamma_\infty\gamma \mid 0<c<K,\ -K^2<d<K^2\right\}\ . \tag{4.5}$$

Here and everywhere else we use (c,d), a shorthand for $(c(\gamma),d(\gamma))$, to denote the lower row entries when writing γ as a 2×2 matrix. Complimented with a term given by the trivial coset (the one with a representative $\gamma=$Id), this is the range of the sum, denoted by $(\Gamma_\infty\backslash\Gamma)_{<K}$, taken in (4.1). With a slight abuse of notation, we use $\gamma\in\Gamma_\infty\backslash\Gamma$ to denote a sum over a representative from each coset.

To ensure that (4.1) is well-defined we require to check that the summands are invariant under the replacement of γ with $\pm T^n\gamma$ for $n\in\mathbb{Z}$. Obviously $\mathrm{reg}(T^n\gamma,\tau)=\mathrm{reg}(\gamma,\tau)$ for any n, and γ and $-\gamma$ act in the same way on $\mathbb{H}$ as well as on the cusp representatives so the Rademacher regularisation factor has the required invariance. Left multiplication by T does not change the lower row of γ so the factor $\mathrm{jac}(\gamma,\tau)^{1/4}$ is invariant under the substitution of $T^n\gamma$ for γ. Using the explicit description (A.6) for ϵ we see that $\epsilon(T^n\gamma)=\mathrm{e}(-\frac{n}{24})\,\epsilon(\gamma)$, so that the product $\epsilon(\gamma)^{-3}\,\mathrm{e}(-\frac{\gamma\tau}{8})$, is also invariant under the substitution of $T^n\gamma$ for γ, and then the identity $\epsilon(-\gamma)^{-3}\,\mathrm{jac}(-\gamma,\tau)^{1/4}=\epsilon(\gamma)^{-3}\,\mathrm{jac}(\gamma,\tau)^{1/4}$ completes the verification that every summand of (4.1) is invariant under the replacement of γ with $\pm T^n\gamma$.

---

generalise the analysis to $\mathrm{SL}_2(\mathbb{R})$ subgroups commensurable to $\mathrm{SL}_2(\mathbb{Z})$

While the regularisation factor might seem ad hoc at the first sight, here we would like to argue that it is in fact strictly necessary and very natural. First, to motivate the appearance of a regularisation note that a series

$$\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \psi(\gamma)\,\mathrm{e}(\alpha\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{w/2}, \tag{4.6}$$

for α a real constant and ψ a (compatible) multiplier system of weight w on Γ, will generally not converge (absolutely and uniformly on compacta) unless $w>2$. (In case it does converge it defines a modular form of weight w on Γ.) In [11] Rademacher demonstrated a method to regularise this sum for $\psi\equiv 1$, $\alpha=-1$ and $w=0$ in the case that $\Gamma=\mathrm{SL}_2(\mathbb{Z})$. His prescription may be described as follows: multiplying each summand by the regularisation factor $\mathrm{r}^0(-1,\gamma,\tau)$ which is 1 when γ is upper triangular and

$$\mathrm{e}(\gamma\tau-\gamma\infty)\,\mathrm{e}(\gamma\infty-\gamma\tau,1)=\mathrm{e}(\gamma\tau-\gamma\infty)(\mathrm{e}(\gamma\infty-\gamma\tau)-1)$$

otherwise, and replacing the sum over the coset space $\Gamma_\infty\backslash\Gamma$ with a limit of sums over rectangles $(\Gamma_\infty\backslash\Gamma)_{<K}$, we obtain

$$\mathrm{e}(-\tau)+\lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} \mathrm{e}(-\gamma\tau)-\mathrm{e}(-\gamma\infty)\ .$$

He then went on to prove (cf. loc. cit.) that this expression converges and defines an $\mathrm{SL}_2(\mathbb{Z})$ invariant function on the upper half-plane, which is nothing but the familiar j-function (up to an additive constant). More generally, in [10], for example, it is shown that for $\psi\equiv 1$ and $\alpha, w/2\in\mathbb{Z}$, inclusion of the factor

$$\mathrm{r}^{w/2}(\alpha,\gamma,\tau)=\mathrm{e}(\alpha(\gamma\infty-\gamma\tau))\,\mathrm{e}(\alpha(\gamma\tau-\gamma\infty),1-w)$$

for summands corresponding to cosets with upper triangular representatives, together with a limit of sums over rectangles, regularises the sum (4.6) for Γ a subgroup of $\mathrm{SL}_2(\mathbb{R})$ that is commensurable with the modular group. Putting $\alpha=-1/8$ and $w=1/2$, we recover exactly our regularisation factor (4.2). It is rather surprising that this straightforward generalisation of regularisation scheme does its job in regularising the sum in (4.1).

To understand the modular properties of R_Γ, it is useful to consider the companion function S_Γ defined by

$$S_\Gamma(\tau)=\lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \epsilon(\gamma)^3\,\mathrm{e}(\tfrac{\gamma\tau}{8})\,\mathrm{jac}(\gamma,\tau)^{3/4}. \tag{4.7}$$

According to the philosophy of the previous paragraph S_Γ, supposing it converges, should have modular transformations of weight 3/2 for Γ with multiplier system coinciding with that of $\eta(\tau)^3$. Also, we should have that $S_\Gamma=\mathcal{O}(q^{1/8})$ as $\tau\to\infty$. It will develop in §7 that R_Γ is a mock modular form and that S_Γ is its shadow.

Remark. The functions $\hat\Sigma(\tau)$ and $\hat\Sigma^\circ(\tau)$ of [20] are recovered by taking $\Gamma=\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma=\Gamma_0(2)$, respectively, in $\hat R_\Gamma(\tau)$, up to a factor of -2. We will demonstrate presently that $R_\Gamma(\tau)=\Sigma(\tau)/2$ for $\Sigma(\tau)$ is as in loc. cit. when $\Gamma=\Gamma_0(1)$.

Apart from considering the Rademacher with $\psi(\gamma)=\epsilon^{-3}(\gamma)$, to construct the M_{24} McKay–Thompson series we also need to extend the above construction to the cases with a more general multiplier system ψ. In particular, we would sometimes like to consider different Rademacher sums with the same group Γ but different multipliers.

Given a group morphism $\rho : \Gamma \to \mathbb{C}^*$ that is trivial on Γ_∞ we may consider the following twists of R_Γ and S_Γ:

$$\begin{aligned} R_{\Gamma,\rho}(\tau) &= \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \psi(\gamma)\, \mathrm{e}(-\tfrac{\gamma\tau}{8})\, \mathrm{reg}(\gamma,\tau)\, \mathrm{jac}(\gamma,\tau)^{1/4} \\ S_{\Gamma,\rho}(\tau) &= \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \bar{\psi}(\gamma)\, \mathrm{e}(\tfrac{\gamma\tau}{8})\, \mathrm{jac}(\gamma,\tau)^{3/4} \end{aligned} \tag{4.8}$$

where we have $\psi(\gamma) = \rho(\gamma)\epsilon(\gamma)^{-3}$. In this paper we will take $\Gamma = \Gamma_0(n)$ for some $n \in \mathbb{Z}^+$ and $\rho = \rho_{n|h}$ given in (2.5) for h simultaneously dividing n and 24. Not surprisingly, later on we are going to take (n,h) to be the (n_g,h_g) coming from the cycle shapes of M_{24} as given by Table 1.

To ease notation, we write $R_{n|h}$ for $R_{\Gamma,\rho}$ when $\Gamma = \Gamma_0(n)$ and $\rho = \rho_{n|h}$, and we apply the directly analogous interpretation to the $S_{n|h}$. The main result in this article is the following.

Theorem 4.1. *Let $g \in M_{24}$. Then $H_g = -2R_{n|h}$ when $n = n_g$ and $h = h_g$.*

The rest of the paper is devoted to the proof of the above theorem.

5 Convergence

After defining the Rademacher sums formally in the last section, in this section we would like to discuss the convergence of the sums. An integral step towards Theorem 4.1 is to prove its convergence. It is a rather subtle issue and in this section we will establish it conditional upon a hypothesis whose validity will be established in §9. To be more precise, the aim in this section is to demonstrate that if the Selberg–Kloosterman zeta function (5.7) attached to $\psi = \rho\epsilon^{-3}$ and $\Gamma = \Gamma_0(n)$ converges at $s = 3/4$ for some character ρ on Γ then the limits (4.1) and (4.7) defining the Rademacher sums $R_{\Gamma,\rho}$ and $S_{\Gamma,\rho}$, respectively, define holomorphic functions on the upper half-plane.

Our arguments in this section owe a lot to the work of Niebur in [31]. We cannot apply his results verbatim since he worked with (mock) modular forms of negative weight and ours have positive weight, but many of the procedures we employ have a counterpart in loc. cit.

Let $\Gamma = \Gamma_0(n)$ for some positive integer n and $\psi(\gamma) = \rho(\gamma)\epsilon(\gamma)^{-3}$ for some character ρ on Γ satisfying $\rho(\gamma) = \rho(T\gamma) = \rho(\gamma T)$. According to the definitions of $R_{\Gamma,\rho}$ and $S_{\Gamma,\rho}$, we can separate the contribution from the trivial coset and the rest in the following way

$$\begin{aligned} R_{\Gamma,\rho}(\tau) &= q^{-1/8} + \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} R_\gamma(\tau) \\ S_{\Gamma,\rho}(\tau) &= q^{1/8} + \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} S_\gamma(\tau)\ . \end{aligned}$$

Using the expression of the (generalised) exponentials as infinite sums and isolating the first term in the sum, we obtain $R_\gamma(\tau) = R^0_\gamma(\tau) + R^+_\gamma(\tau)$ and $S_\gamma(\tau) = S^0_\gamma(\tau) + S^+_\gamma(\tau)$, where

$$\begin{aligned} R^0_\gamma &= e(\tfrac{1}{8})\, \psi(\gamma)\, \mathrm{e}(-\tfrac{\gamma\infty}{8})\, c^{-3/2}\, (\tau - \gamma^{-1}\infty)^{-1} \\ S^0_\gamma &= \bar{\psi}(\gamma)\, \mathrm{e}(\tfrac{\gamma\infty}{8})\, c^{-3/2}\, (\tau - \gamma^{-1}\infty)^{-3/2} \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} R^+_\gamma &= \psi(\gamma)\,\mathrm{e}(-\tfrac{\gamma\infty}{8})\sum_{m=1}^{\infty}\frac{c^{-2m-3/2}(\tau-\gamma^{-1}\infty)^{-m-1}(2\pi i/8)^{m+1/2}}{\Gamma(m+3/2)} \\ S^+_\gamma &= \bar{\psi}(\gamma)\,\mathrm{e}(\tfrac{\gamma\infty}{8})\sum_{m=1}^{\infty}\frac{c^{-2m-3/2}(\tau-\gamma^{-1}\infty)^{-m-3/2}(-2\pi i/8)^{m}}{m!}\ . \end{aligned} \tag{5.2}$$

We have also applied the identities

$$\mathrm{jac}(\gamma,\tau)=c^{-2}(\tau-\gamma^{-1}\infty)^{-2}\quad,\quad \gamma\infty-\gamma\tau=c^{-2}(\tau-\gamma^{-1}\infty)^{-1}$$

in rewriting the above expressions. With these definitions, we have the following:

Lemma 5.1. *The sums* $\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times}R^+_\gamma(\tau)$ *and* $\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times}S^+_\gamma(\tau)$ *are absolutely convergent, locally uniformly for* $\tau\in\mathbb{H}$.

Proof. We give details for the first sum. A very similar argument takes care of the second one.

By (5.2) we have that $R^+_\gamma(\tau)=\mathcal{O}(c^{-7/2}\,|\tau-\gamma^{-1}\infty|^{-2})$ as $c\to\infty$. For $\Gamma=\Gamma_0(n)$, the representatives of the non-trivial cosets in the coset space $\Gamma_\infty\backslash\Gamma$ are indexed by their lower rows which are tuples of two integers (c,d) with $c>0$ and $n|c$, $\gcd(c,d)=1$. A procedure we will repeatedly use is to split the sum over d into the following double sum

$$\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times}f(c,d)=\sum_{\substack{0<c\\ n|c}}\sum_{\substack{0\le d'<c\\ (c,d')=1}}\sum_{\ell\in\mathbb{Z}}f(c,d'+\ell c) \tag{5.3}$$

for an arbitrary function $f(x,y)$. Notice that the above sum over the tuple (c,d') can be viewed as the sum over a complete and irredundant set of representatives of a double coset in $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$.

Taking $f(x,y)=x^{-7/2}|\tau+y/x|^{-2}$ we find that

$$\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times}|R^+_\gamma(\tau)|\ll\sum_{\substack{0<c\\ n|c}}c^{-7/2}\sum_{\substack{0\le d'<c\\ (c,d')=1}}\sum_{\ell\in\mathbb{Z}}|\tau+d'/c+\ell|^{-2}.$$

Since there are not more than c invertible elements in the ring $\mathbb{Z}/c$, we conclude that the sum $\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times}R^+_\gamma(\tau)$ is absolutely convergent, locally uniformly for $\tau\in\mathbb{H}$, as required. ☐

Hence, in order to show the convergence of the sums defining $R_{\Gamma,\rho}(\tau)$ and $S_{\Gamma,\rho}(\tau)$ it now suffices to establish the convergence of the limits

$$\lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}}R^0_\gamma(\tau),\quad \lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}}S^0_\gamma(\tau), \tag{5.4}$$

and apply Lemma 5.1.

After carefully examining the limit and employing the Lipschitz summation formula in the form of Lemma C.1, the first of the above sums can be rewritten as

$$\lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}}R^0_\gamma(\tau)=2\pi\,\mathrm{e}(-\tfrac18)\sum_{k=1}^{\infty}\mathrm{e}((k-\tfrac18)\tau)\lim_{K\to\infty}\sum_{\substack{0<c<K\\ n|c}}c^{-3/2}S(0,k,c,\psi)\ , \tag{5.5}$$

where we have used the *generalised Kloosterman sum* (cf. [32]) for a given group Γ:

$$S(m,\ell,c,\xi)=\sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_\infty\\ c(\gamma)=c}}\xi(\gamma)\,\mathrm{e}((m-\alpha)\gamma\infty)\,\mathrm{e}(-(\ell-\alpha)\gamma^{-1}\infty) \tag{5.6}$$

for $\ell\in\mathbb{Z}$, where for $m,\ell,c\in\mathbb{Z}$ with $0<c$ and ξ a multiplier system on Γ, the value $0\leq\alpha<1$ is determined by ξ according to the requirement that $\xi(T)=\mathrm{e}(\alpha)$. An important object closely related to the generalised Kloosterman sum is the *Selberg–Kloosterman zeta function* of Γ (cf. loc. cit.), defined as

$$Z_{m,\ell}(s,\psi)=\sum_{c>0}\frac{S(m,\ell,c,\psi)}{c^{2s}}\,. \tag{5.7}$$

The details of the manipulation leading to (5.5) are given in Appendix D.

Hence, as promised, if we assume the convergence of $Z_{0,k}(s,\rho\epsilon^{-3})$ at $s=3/4$ we will have shown the convergence of the first sum in (5.4), and hence $R_{\Gamma,\rho}$ upon using Lemma (5.1). The convergence of the Selberg–Kloosterman zeta function will be demonstrated in §9. Finally, a similar but less subtle argument also establishes the convergence of the second sum in (5.4). In particular, in the case of S^0_γ the application of the Lipschitz summation formula is less delicate: We can apply (C.1) with $s=3/2,\alpha=7/8$ and the error term E_K is absent. Taken together we have the following:

Proposition 5.2. *The expression (4.1) defining $R_{\Gamma,\rho}(\tau)$ converges, locally uniformly for $\tau\in\mathbb{H}$, thus defining a holomorphic function on $\mathbb{H}$. Also, the expression (4.7) defining $S_{\Gamma,\rho}(\tau)$ converges, locally uniformly for $\tau\in\mathbb{H}$, thus defining a holomorphic function on $\mathbb{H}$.*

In practice it is useful to know that there is some flexibility in the sum over “rectangles” in the definition of $R_{\Gamma,\rho}$ and $S_{\Gamma,\rho}$ (cf. (4.1) and (4.7)). For example, given $\sigma\in\Gamma$ we may consider the sums

$$\lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma}R_\gamma(\tau),\qquad \lim_{K\to\infty}\sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma}S_\gamma(\tau), \tag{5.8}$$

where $(\Gamma_\infty\backslash\Gamma)_{<K}\sigma$ denotes the “parallelogram” that is the image of $(\Gamma_\infty\backslash\Gamma)_{<K}$ under right multiplication by σ. The following result can be proved by using the technique employed in Lemma 4.3 of [31] to extend the convergence arguments appearing already in this section. We suppress the details.

Lemma 5.3. *Let $\sigma\in\Gamma$, and let $R_\gamma(\tau)$ and $S_\gamma(\tau)$ be as in (5.1) and (5.2), respectively. Then the expressions in (5.8) converge uniformly for $\tau\in\mathbb{H}$, thus defining holomorphic functions on $\mathbb{H}$. Moreover these functions coincide with $R_{\Gamma,\rho}(\tau)$ and $S_{\Gamma,\rho}(\tau)$ respectively.*

6 Coefficients

A common motivation to study Rademacher sums is to obtain an exact expression for the Fourier coefficients of a given infinite q-series. In this section we will give such expressions for the functions $R_{\Gamma,\rho}$ and $S_{\Gamma,\rho}$. They are useful, for instance, in the physical study of the growth (6.2) of the twisted indices of supersymmetric states—the (twisted) entropy—as the energy of the states gets large.

In the last section we have seen that, by splitting the sum over representatives of cosets $\Gamma_\infty\backslash\Gamma$ into a double sum over those of the double cosets $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$ and over all integers $\ell\in\mathbb{Z}$, and applying the Lipschitz summation formula (C.2), we obtain expressions (5.5) for the objects (5.4) as an infinite q-series. In this section we would like to show that the same can be done for the full $R_{\Gamma,\rho}, S_{\Gamma,\rho}$, and write down explicit expressions for the infinite q-series that the infinite sums (4.8) converge to.

Following similar steps as delineated in the last section, this time with the help of equation (C.1), it is not difficult to prove that the functions $R_{\Gamma,\rho}, S_{\Gamma,\rho}$ can be written as q-series with the following Fourier coefficients

$$\begin{aligned} c_{\Gamma,\rho}(k-\tfrac{1}{8}) &= \frac{2\pi\ \mathrm{e}(-\frac{1}{8})}{(8k-1)^{1/4}} \sum_{\substack{0<c\\ n|c}} \frac{1}{c} I_{1/2}\left(\frac{\pi}{2c}(8k-1)^{1/2}\right) S(0,k,c,\psi) \\ c^{*}_{\Gamma,\rho}(k+\tfrac{1}{8}) &= \mathrm{e}(-\tfrac{3}{8})\, 2\pi (8k+1)^{1/4} \sum_{\substack{0<c\\ n|c}} \frac{1}{c} J_{1/2}\left(\frac{\pi}{2c}(8k+1)^{1/2}\right) S(1,k+1,c,\bar{\psi})\ , \end{aligned} \tag{6.1}$$

where $S(m,\ell,c,\psi)$ again denotes the generalised Kloosterman sum (5.6) and $\psi = \rho\epsilon^{-3}$.

First note that the q-series with the above coefficients indeed define holomorphic functions on $\mathbb{H}$. To start with, again assuming the convergence of the Selberg–Kloosterman zeta function $Z_{0,k}(\frac{3}{4},\psi)$, we will show that the sums (over c) in (6.1) converge: Observe that since $c^{-1}I_{1/2}(\frac{\pi}{2c}\sqrt{8k-1})$ is bounded by a constant times $c^{-3/2}$ for c sufficiently large (cf. (B.5)), the convergence of the right hand side of (6.1) is guaranteed by the convergence of $Z_{0,k}(\frac{3}{4},\psi)$. We establish the convergence of $Z_{0,k}(\frac{3}{4},\psi)$ for $k\in\mathbb{Z}$ in §9 for the case that $\Gamma=\Gamma_0(n)$ and $\rho=\rho_{n|h}$ for some h dividing both n and 24. From the above argument we see that this result implies the convergence of both the $c_{\Gamma,\rho}(k-\frac{1}{8})$ for $k>0$ and the $c^{*}_{\Gamma,\rho}(k+\frac{1}{8})$ for $k\geq 0$.

Once armed with the convergence of the terms $c_{\Gamma,\rho}(k-\frac{1}{8})$ and $c^{*}_{\Gamma,\rho}(k+\frac{1}{8})$ we may consider the following series

$$\sum_{k\geq 1} c_{\Gamma,\rho}(k-\tfrac{1}{8})q^{k-1/8}, \quad \sum_{k\geq 0} c^{*}_{\Gamma,\rho}(k+\tfrac{1}{8})q^{k+1/8}.$$

The Bessel function $I_{1/2}(x)$ is asymptotic to $e^x/\sqrt{2\pi x}$ for large x (cf. §B) so $c_{\Gamma,\rho}(k-\frac{1}{8})$ is dominated by the first term in the summation over c in (6.1) for n sufficiently large, and similarly for $c^{*}_{\Gamma,\rho}(k+\frac{1}{8})$; in the case that $\Gamma=\Gamma_0(n)$ this is the term with $c=n$. We then have

$$c_{\Gamma,\rho}(k-\tfrac{1}{8}) = \mathcal{O}\left(\frac{e^{\pi\sqrt{8k-1}/2n}}{\sqrt{8k-1}}\right) \tag{6.2}$$

as $k\to\infty$ and a similar estimate holds for $c^{*}_{\Gamma,\rho}(k+\frac{1}{8})$. In particular, we may conclude that the series $\Phi_{\Gamma,\rho}$ and $\Phi^{*}_{\Gamma,\rho}$, defined by

$$\Phi_{\Gamma,\rho}(\tau) = q^{-1/8} + \sum_{k\geq 1} c_{\Gamma,\rho}(k-\tfrac{1}{8})q^{k-1/8}, \tag{6.3}$$

$$\Phi^{*}_{\Gamma,\rho}(\tau) = q^{1/8} + \sum_{k\geq 0} c^{*}_{\Gamma,\rho}(k+\tfrac{1}{8})q^{k+1/8}, \tag{6.4}$$

converge absolutely and locally uniformly for $\tau\in\mathbb{H}$ upon identifying $q=\mathrm{e}(\tau)$.

We summarise the discussion of the previous two paragraphs in the following lemma.

Lemma 6.1. *The expressions (6.1) defining* $c_{\Gamma,\rho}(k-\frac{1}{8})$ *and* $c^*_{\Gamma,\rho}(k+\frac{1}{8})$ *converge. Further, the generating series (6.3) and (6.4) for the* $c_{\Gamma,\rho}(k-\frac{1}{8})$ *and* $c^*_{\Gamma,\rho}(k+\frac{1}{8})$ *converge absolutely and locally uniformly for* $\tau \in \mathbb{H}$*, thus defining holomorphic functions* $\Phi_{\Gamma,\rho}(\tau)$ *and* $\Phi^*_{\Gamma,\rho}(\tau)$ *on* $\mathbb{H}$*. These functions coincide with* $R_{\Gamma,\rho}(\tau)$ *and* $S_{\Gamma,\rho}(\tau)$ *respectively.*

7 Variance

After establishing the convergence of the Rademachers sums $R_{\Gamma,\rho}$ and $S_{\Gamma,\rho}$, in this section we wish to determine how they transform under the action of Γ. In particular, we will show that $S_{\Gamma,\rho}$ is a weight 3/2 cusp form and $R_{\Gamma,\rho}$ is a weight 1/2 mock modular form on Γ. Moreover, the shadow of $R_{\Gamma,\rho}$ is given by $S_{\Gamma,\rho}$.

As before, we have $\Gamma = \Gamma_0(n)$ for some positive integer n, and let $\rho = \rho_{n|h}$ (cf. (2.5)) for some h dividing both n and 24 and set $\psi = \rho\epsilon^{-3}$. The variance of the Rademacher sum $S_{\Gamma,\rho}$ can be easily established, and we have

Proposition 7.1. *The function* $S_{\Gamma,\rho}(\tau)$ *is a cusp form of weight* $3/2$ *for* Γ *with multiplier system* $\overline{\psi}$*.*

Proof. Since ϵ^3 is the multiplier of the cusp form $\eta^3(\tau)$ and ρ is a group character satisfying $\rho(\gamma)\rho(\sigma) = \rho(\gamma\sigma)$ for all $\gamma, \sigma \in \Gamma_0(n)$, we see that $\bar{\psi} = \rho^{-1}\epsilon^3$ is a weight 3/2 multiplier system, which leads immediately to

$$S_\gamma|_{\overline{\psi},3/2}\sigma = S_{\gamma\sigma} \tag{7.1}$$

for $\sigma \in \Gamma$, where S_γ is as defined in (5.1)-(5.2). Moreover, when $h = 1$ and hence ρ is trivial, this also holds more generally for $\sigma \in \mathrm{SL}_2(\mathbb{Z})$.

Applying Lemma 5.3 to $S_{\Gamma,\rho} = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} S_\gamma(\tau)$ and using (7.1), we immediately have $S_{\Gamma,\rho}|_{\overline{\psi},3/2}\sigma = S_{\Gamma,\rho}$ for $\sigma \in \Gamma$. This establishes that $S_{\Gamma,\rho}$ is a modular form of weight 3/2 for Γ with multiplier system $\overline{\psi}$. We also need to show that $S_{\Gamma,\rho}$ vanishes at all cusps. For the cusp at ∞, this is clear from Lemma 6.1. We will now show that it is also true for the other cusps of Γ.

An important property of a given cusp is its width. Suppose $\kappa \in \mathbb{Q}$ is a representative of a cusp of $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ that is related to the infinite cusp by $\sigma\infty = \kappa$, $\sigma \in \mathrm{SL}_2(\mathbb{Z})$. We define the *width* u of the cusp κ to be the positive integer such that the subgroup of Γ stabilising κ is given by

$$\Big\{\gamma \in \Gamma \Big| \gamma\kappa = \kappa\Big\} = \sigma\langle \pm T^u\rangle\sigma^{-1} = \sigma\Gamma^u_\infty\sigma^{-1}\ . \tag{7.2}$$

Alternatively, we can rephrase the above definition using the concept of a 'scaling element' which will become important later. Consider an element $\tilde{\sigma} \in \mathrm{SL}_2(\mathbb{R})$, to be called a *scaling element* of Γ at cusp κ, such that $\kappa = \tilde{\sigma}\infty$ and the subgroup of Γ stabilising κ is given by $\tilde{\sigma}\Gamma_\infty\tilde{\sigma}^{-1}$ (Cf. [10, §2.6]). Such a scaling element necessarily has the form

$$\tilde{\sigma} = \sigma U T^\beta \quad , \quad \text{with} \quad \sigma \in \mathrm{SL}_2(\mathbb{Z})\ ,\ \beta \in \mathbb{Q}\ , \tag{7.3}$$

where $U = \left(\begin{smallmatrix}\sqrt{u} & 0\\ 0 & 1/\sqrt{u}\end{smallmatrix}\right)$ maps $\tau \to u\tau$ and $T^\beta = \left(\begin{smallmatrix}1 & \beta\\ 0 & 1\end{smallmatrix}\right)$.

Given a cusp representative $\kappa = \sigma\infty$ of width v of $\Gamma_0(n)$, in order to study the behaviour of $S_{n|h}$ at κ, it is convenient to consider the function $S^{(\kappa)}_{n|h}(\tau)$ which we will now define. For a given scaling element $\tilde{\sigma}$, for $h = 1$ we define

$$S^{(\kappa)}_{n|1}(\tau) = S_{n|1}\big|_{\epsilon^3,3/2}\sigma(v(\tau+\beta)) = \epsilon^3(\sigma)\,\mathrm{jac}(\sigma, v(\tau+\beta))^{3/4} S_{n|h}(\tilde{\sigma}\tau)\ .$$

From (7.1) we get

$$S^{(\kappa)}_{n|1}(\tau) = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} S_\gamma(v(\tau+\beta)) \ .$$

Splitting the above sum into a sum over the double coset $(\Gamma_\infty\backslash\Gamma)_{<K}\sigma/\Gamma^v_\infty$ and applying the Lipschitz summation formula as before, we obtain that $S^{(\kappa)}_{n|1}(\tau)$ takes the form of an infinite series $\sum_{k=0}^\infty c^{*(\kappa)}_{\Gamma,\rho}(k+\nu)\,\mathrm{e}((k+\nu)\tau)$ where $0<\nu\le 1$ differs from $\frac{v}{8}$ by an integer. This proves that $S_{n|1}$ vanishes at κ.

More generally, for $h\neq 1$ we again consider

$$\begin{aligned} S_{n|h}\big|_{\epsilon^3,3/2}\sigma(\tau) &= \epsilon^3(\sigma)\,\mathrm{jac}(\sigma,\tau)^{3/4} S_{n|h}(\sigma\tau) \\ &= \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} \rho^{-1}(\gamma\sigma^{-1})\epsilon^3(\gamma)\,\mathrm{jac}(\gamma,\tau)^{3/4}\,\mathrm{e}(\tfrac{\gamma\tau}{8}) \ . \end{aligned} \tag{7.4}$$

If κ has width v in $\Gamma_0(n)$ and has width u in $\Gamma_0(nh)$, we can define for a given scaling element $\tilde\sigma$

$$S^{(\kappa)}_{n|h}(\tau) = S_{n|h}\big|_{\epsilon^3,3/2}\sigma\,(u(\tau+\beta)) \ .$$

To obtain its expression as an infinite series, define $0\le m<h$ such that $\rho(\sigma T^v\sigma^{-1}) = \mathrm{e}(\frac{m}{h})$ and proceed as before, ultimately obtainin the general expression

$$S^{(\kappa)}_{n|h}(\tau) = \sum_{k=0}^\infty c^{*(\kappa)}_{\Gamma,\rho}(k+\nu)\,\mathrm{e}((k+\nu)\tfrac{u}{v}\tau) \quad , \quad 0<\nu\le 1 \ . \tag{7.5}$$

More precisely, $\nu = \{\{\frac{m}{h}+\frac{v}{8}\}\}$ is given by the sawtooth function $\{\{x\}\}$ defined as the difference between a real number and the largest integer that is smaller than it:

$$\{\{x\}\} = \begin{cases} x-\lfloor x\rfloor & x\in\mathbb{R}\backslash\mathbb{Z} \\ 1 & x\in\mathbb{Z} \end{cases}$$

where the floor function $\lfloor x\rfloor$ denotes the largest integer not exceeding x. This shows that $S_{n|h}(\tau)$ vanishes at all cusps of $\Gamma_0(n)$ and thereby finishes the proof. ☐

Our next task is to establish the variance of the Rademacher sum $R_{\Gamma,\rho}$.

Proposition 7.2. *The function $R_{\Gamma,\rho}$ is a mock modular form of weight $1/2$ on Γ with multiplier system ψ and with shadow $S_{\Gamma,\rho}$.*

Proof. First observe that $R_\gamma = \psi(\gamma)\,\mathrm{e}(-\frac{\gamma\tau}{8})\,\mathrm{jac}(\gamma,\tau)^{1/4}\mathrm{reg}(\gamma,\tau)$ would transform nicely if it weren't for the regularisation factor. Therefore, it would be useful to separate from R_γ the "unregularised" part

$$P_\gamma(\tau) = \psi(\gamma)\,\mathrm{e}(-\gamma\tau/8)\,\mathrm{jac}(\gamma,\tau)^{1/4}, \tag{7.6}$$

for $\gamma\in\Gamma$, and consider

$$R_\gamma(\tau) - P_\gamma(\tau) = \psi(\gamma)\,\mathrm{e}(-\tfrac{\gamma\infty}{8})\,\mathrm{jac}(\gamma,\tau)^{1/4}\,(\mathrm{e}(X,1/2)-\mathrm{e}(X))$$

with $X = \frac{\gamma\infty - \gamma\tau}{8}$. Using the definition (4.4), we have

$$\mathrm{e}(X,s) - \mathrm{e}(X) = -\frac{\mathrm{e}(X)}{\Gamma(s)} \int_{2\pi i X}^{\infty} e^{-t} t^{s-1}\, dt \;.$$

Setting $t = \frac{2\pi i}{8}(\gamma z - \gamma\tau)$ in the integral and using the identity

$$\mathrm{jac}(\gamma, z)\,\mathrm{jac}(\gamma,\tau)(z-\tau)^2 = (\gamma z - \gamma\tau)^2 \tag{7.7}$$

we get

$$\mathrm{e}(X,1/2) - \mathrm{e}(X) = \frac{1}{2}\,\mathrm{e}(\tfrac{1}{8})\,\mathrm{e}(\tfrac{\gamma\infty}{8})\,\mathrm{jac}(\gamma,\tau)^{-1/4} \int_{\gamma^{-1}\infty}^{\infty} \mathrm{e}(-\tfrac{\gamma z}{8})\,\mathrm{jac}^{3/4}(\gamma,z)\,(z-\tau)^{-1/2}\, dz \;.$$

It will be convenient to introduce an integral operator J_x for x in the extended upper half-plane $(\mathbb{H} \cup \mathbb{Q} \cup \infty)$ which acts on a holomorphic function g on $\mathbb{H}$ (with sufficiently rapid decay as $\Im(z) \to \infty$) as

$$(J_x g)\,(\tau) = \frac{\mathrm{e}(\frac{1}{8})}{2} \int_{\bar{x}}^{\infty} \bar{g}(z)(z-\tau)^{-1/2} \mathrm{d}z \;. \tag{7.8}$$

From the definition (3.4), we see that we have a weight 1/2 mock modular form with shadow g if $\hat{h} = h + J_\tau g$ transforms as a weight 1/2 modular form.

Comparing with the definition of the integral operator J_x, we have

$$R_\gamma(\tau) - P_\gamma(\tau) = \left(J_{\gamma^{-1}\infty} S_\gamma\right)(\tau) \;.$$

Note in particular that this is trivially true when γ is in the trivial coset given by upper triangular matrices, since both sides of the equation vanish.

By the fact that ψ is a multiplier system for Γ of weight 1/2, we immediately have for an element σ of Γ

$$P_\gamma|_{\psi,1/2}\sigma = P_{\gamma\sigma} \;, \text{ and } \; (J_{\gamma^{-1}\infty} S_\gamma)|_{\psi,1/2}\sigma = (J_{\sigma^{-1}\gamma^{-1}\infty} - J_{\sigma^{-1}\infty}) S_{\gamma\sigma} \;\;.$$

Upon summation over the representatives of the rectangle $(\Gamma_\infty\backslash\Gamma)_{<K}, K \to \infty$ and evoking Lemma 5.3, we obtain

$$R_{\Gamma,\rho}|_{\psi,1/2}\,\sigma = R_{\Gamma,\rho} - J_{\sigma^{-1}\infty} S_{\Gamma,\rho} \;. \tag{7.9}$$

Now it is an easy calculation to show that[2]

$$\left(R_{\Gamma,r} + J_\tau S_{\Gamma,\rho}\right)\Big|_{\psi,1/2}\sigma = R_{\Gamma,r} + J_\tau S_{\Gamma,\rho} \;,$$

which by definition (3.4) proves the claim. ☐

---

[2] In fact, one can show that the non-holomorphic completion $R_\Gamma + \tilde{S}_\Gamma$ also has a Rademacher form given by

$$\hat{R}_\Gamma(\tau) = \lim_{K\to\infty} \sum_{\gamma \in (\Gamma_\infty\backslash\Gamma)_{<K}} \epsilon(\gamma)^{-3}\,\mathrm{e}(-\tfrac{\gamma\tau}{8})\,\mathrm{com}(\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{1/4} \tag{7.10}$$

where $\mathrm{com}(\tau)$ is the *non-holomorphic completion factor* given in terms of the incomplete Gamma function, $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^t \mathrm{d}t$, by

$$\mathrm{com}(\tau) = 1 - \Gamma(1/2, \pi\Im(\tau)/2)/\sqrt{\pi} \;.$$

8 Coincidence

Recall from §4 that we write $R_{n|h}$ for $R_{\Gamma,\rho}$ when $\Gamma = \Gamma_0(n)$ and $\rho = \rho_{n|h}$, and similarly for $S_{n|h}$. In §7 we have established the transformation property of $R_{n|h}$ and $S_{n|h}$ under the group $\Gamma_0(n)$. In this section we will explicitly identify them. More precisely, we will first prove in §8.1 that

$$S_{n|h} = \begin{cases} \lambda_n \eta^3, \lambda_n \in \mathbb{C} & h = 1 \\ 0 & h > 1 \end{cases} \tag{8.1}$$

for the $(n,h) = (n_g, h_g)$ that appear in the cycle shapes of $[g] \subset M_{24}$, (cf. Table 1), and later in §8.2 complete the proof our main theorem 4.1 by showing $H_g = -2R_{n|h}$, assuming the result of §9.

8.1 Determining S

Let n and h be positive integers such that h divides both n and 24 and consider the function

$$F_{n|h}(\tau) = \frac{S_{n|h}(\tau)}{\eta(\tau)^3} \; .$$

Since $S_{n|h}$ has the same multiplier system as η^3 on $\Gamma_0(nh)$ and η does not vanish on the upper half-plane, $F_{n|h}$ is a modular function for $\Gamma_0(nh)$ that satisfies the transformation rule

$$F_{n|h}(\gamma\tau) = \rho(\gamma) F_{n|h}(\tau) \tag{8.2}$$

for $\gamma \in \Gamma_0(n)$ where $\rho = \rho_{n|h}$ is given explicitly by (2.5). In particular, $F_{n|h}$ induces a morphism, which we denote by $f_{n|h}$, of Riemann surfaces $X_0(nh) \to \mathbb{P}^1$ where $X_0(N)$ denotes the *modular curve* of level N and is formed by taking the quotient of the *extended upper half-plane*

$$\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$$

by the action of the group $\Gamma_0(N)$. We will denote the corresponding canonical map by

$$\varphi_N : \mathbb{H}^* \to X_0(N) \; . \tag{8.3}$$

Note that there are no non-constant maps $X_0(nh) \to \mathbb{C}$, so either $F_{n|h}$ has a pole at some cusp of $\Gamma_0(nh)$, or $S_{n|h} = \lambda_{n|h}\eta^3$ for some $\lambda_{n|h} \in \mathbb{C}$. Of course, $S_{n|h}$ and η^3 have different automorphy for $\Gamma_0(n)$ when $h > 1$, and an identity $S_{n|h} = \lambda_{n|h}\eta^3$ implies that $\lambda_{n|h}$ and hence $S_{n|h}$ are identically zero. Our main strategy to show (8.1) is to show that $F_{n|h}$ is bounded at all cusps of $\Gamma_0(nh)$ for all the pairs (n,h) under consideration.

Since the Fourier expansion of both $S_{n|h}$ and η^3 around the infinite cusp starts with a term $q^{1/8}$(cf. (6.1)), we see that $F_{n|h}(\tau)$ is bounded as $\tau \to i\infty$ and can only have poles at the cusps of $\Gamma_0(nh)$ that are not represented by infinity. Furthermore, from the transformation (8.2) of $F_{n|h}$ under $\Gamma_0(n)$, we see that $f_{n|h}$ has a pole at a given cusp of $\Gamma_0(nh)$ if and only if it has a pole of the same order at every cusp that is an image of it under $\Gamma_0(n)$. Therefore, to prove (8.1) it is sufficient to show that $F_{n|h}$ are bounded at all cusps of $\Gamma_0(n)$ other than the infinite cusp.

In order to analyse the behaviour of $F_{n|h}$ at the cusp representative $\kappa \in \mathbb{Q}$, as before we will now consider its transformation under a given scaling element (cf. (7.3))

$$\tilde{\sigma} = \sigma \begin{pmatrix} \sqrt{u} & 0 \\ 0 & 1/\sqrt{u} \end{pmatrix} \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad , \quad \sigma \in \mathrm{SL}_2(\mathbb{Z}) \; , \; \beta \in \mathbb{R}$$

n	1	2	3	4	5	6	7	8	10	11	12	14	15	21	23
cusps	/	0	0	$0,\frac{1}{2}$	0	$0,\frac{1}{2},\frac{1}{3}$	0	$0,\frac{1}{2},\frac{1}{4}$	$0,\frac{1}{2},\frac{1}{5}$	0	$0,\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{6}$	$0,\frac{1}{2},\frac{1}{7}$	$0,\frac{1}{3},\frac{1}{5}$	$0,\frac{1}{3},\frac{1}{7}$	0
widths	/	2	3	4,1	5	6,3,2	7	8,2,1	10,5,2	11	12,3,4,3,1	14,7,2	15,5,3	21,7,3	23

Table 3: In this table we list the cusps other than the one at $i\infty$, which has width 1, of the groups $\Gamma_0(n)$ for $n = n_g$ in Table 1. Note that, in these cases we have an extra (non-infinite) cusp, represented by $\frac{1}{q}$, for every divisor q of n, and its width is given by the smallest positive integer v such that $n|qv^2$. Moreover, the corresponding multiplier $\rho(\sigma T^v \sigma^{-1}) = \mathrm{e}(\frac{m}{h})$ is given by $m = \frac{q^2 v}{n}(1+qv) \bmod h$.

for $\Gamma_0(nh)$ at $\kappa = \tilde{\sigma}\infty$:

$$F^{(\kappa)}_{n|h}(\tau) = F_{n|h}(\tilde{\sigma}\tau) = \frac{S_{n|h}(\tilde{\sigma}\tau)}{\eta^3(\tilde{\sigma}\tau)} = \frac{S^{(\kappa)}_{n|h}(\tau)}{\eta^3(u(\tau+\beta))} \ . \tag{8.4}$$

When $\tilde{\sigma}$ normalises $\Gamma_0(nh)$, $F^{(\kappa)}_{n|h}(\tau)$ again defines a modular function for $\Gamma_0(nh)$. The order of poles or zeros of the function $F^{(\kappa)}_{n|h}(\tau)$ at $\tau \to i\infty$ gives the order of poles or zeros of the function $f_{n|h} : X_0(nh) \to \mathbb{P}^1$ at the point x given by the image of the κ under the canonical map φ_{nh}.

In the proofs, we will be especially interested in the scaling elements that are elements of the so-called Atkin-Lehner involutions W_u, which are defined for every exact divisor u of n as the set of matrices of the form

$$\frac{1}{\sqrt{u}}\begin{pmatrix} au & b \\ cn & du \end{pmatrix} \quad , \quad a,b,c,d \in \mathbb{Z}$$

with determinant 1. We say u is an exact divisor of n if $(n/u, u) = 1$. An important element of W_n is the *Fricke involution*, given by

$$w_n = \begin{pmatrix} 0 & -1/\sqrt{n} \\ \sqrt{n} & 0 \end{pmatrix} \ . \tag{8.5}$$

From the expansion of $S^{(\kappa)}_{n|h}(\tau)$ recorded in (7.5), we see that $F^{(\kappa)}_{n|h}(\tau)$ has the expansion

$$F^{(\kappa)}_{n|h}(\tau) = \sum_{k=0}^{\infty} C_{n|h}(k)\, q^{\ell + \frac{u}{v}k} \quad , \quad \ell = \tfrac{u}{v}\,\{\{\tfrac{m}{h} + \tfrac{v}{8}\}\} - \tfrac{u}{8} \ . \tag{8.6}$$

As before, u denotes the width of the cusp in $\Gamma_0(nh)$, v denotes the width of the cusp in $\Gamma_0(n)$, and $\mathrm{e}(\frac{m}{h}) = \rho(\sigma T^v \sigma^{-1})$ as defined in (7.2). Moreover, since $F^{(\kappa)}_{n|h}(\tau)$ is invariant under $\langle \pm T \rangle \subset \tilde{\sigma}^{-1}\Gamma_0(nh)\tilde{\sigma}$, its Fourier expansion can only involve integral powers of q and it follows immediately that

$$F^{(\kappa)}_{n|h}(\tau) = \mathcal{O}(q^{\ell}) \quad , \quad \ell \geq 1 + \lfloor -\tfrac{u}{8} \rfloor \quad , \quad \ell \in \mathbb{Z} \ . \tag{8.7}$$

More generally, from the above formula it is also clear that, to prove that $f_{n|h}$ is bounded at κ it is sufficient to prove that

$$\{\{\tfrac{m}{h} + \tfrac{v}{8}\}\} \geq \tfrac{v}{8} \ . \tag{8.8}$$

We shall first consider the cases $n \leq 8$ when there are no cusps of $\Gamma_0(n)$ with width larger than 8. For these cases we have the following Lemma:

Lemma 8.1. *Consider the group $\Gamma_0(n)$ with $1 \leq n \leq 8$ and multiplier $\rho_{n|h}$ defined in (2.5) with h dividing both n and 24. The pole-free condition (8.8) is satisfied for all cusps of $\Gamma_0(n)$ if and only if $(n,h) = (n_g, h_g)$ is one of the 13 pairs that correspond to the cycle shapes of M_{24}, as collected in Table 1.*

Proof. First consider the case $h = 1$. In this case, since there are no cusps with $v > 8$, we get from (8.7) that $f_{n|h}$ has no pole for all $1 \leq n \leq 8$, $h = 1$.

For $h > 1$, we need more refined information about the order of zeros of $S^{(\kappa)}_{n|h}$ to show that, for a given $n < 9$, among all the possible values of h that divide both n and 24, the condition (8.8) is satisfied if and only if $(n,h) = (n_g, h_g)$ is one of the 13 pairs that correspond to the cycle shapes of M_{24}.

For the cusp with a representative $1/q$ and width v, taking $\sigma = \left(\begin{smallmatrix} 1 & 1 \\ q & q+1 \end{smallmatrix}\right)$ we get

$$m = \tfrac{q^2 v}{n}(1 + qv) \mod h\ . \tag{8.9}$$

In particular, we have $m = q$ and $v = n/q$ if q is an exact divisor of n. For the zero cusp ($q = 1$, note that 0 and 1 are related by $T \in \Gamma_0(n)$), therefore, we need h to satisfy $\{\{\frac{1}{h} + \frac{n}{8}\}\} \geq \frac{n}{8}$. It is straightforward to check that this condition is satisfied if and only if $(n,h) = (n_g, h_g)$ is one of the 13 pairs that show up in Table 1.

It remains to be shown that this condition is also satisfied at cusps other than 0 or ∞. First we check that $\{\{\frac{q}{h} + \frac{n/q}{8}\}\} \geq \frac{n/q}{8}$ is indeed satisfied for $n = h = 6$, $q = 2$ and $q = 3$. Finally, for $n = 4$, we have $q = 2$, $v = 1$ and $m = 3$ and we need to check $\{\{\frac{3}{h} + \frac{1}{8}\}\} \geq \frac{1}{8}$, which is indeed the case for $h = 4$. ☐

Corollary 8.2. *For $(n,h) = (n_g, h_g)$ being one of the 13 pairs with $1 \leq n \leq 8$ that arise from in the cycle shapes of M_{24}, (cf. Table 1), we have $S_{n|1} = \lambda_n \eta^3$ and $S_{n|h} = 0$ for $h > 1$.*

For $n > 8$, a direct proof via (8.8) is less straightforward and instead we would like to use arguments involving the topology of $X_0(n)$. When $X_0(nh)$ has genus one, we have the following two lemmas:

Lemma 8.3. *If $n \in \{11, 14, 15\}$ then $S_{n|1} = \lambda_n \eta^3$ for some $\lambda_n \in \mathbb{C}$.*

Proof. For these values of n there is just one cusp of $\Gamma_0(n)$ represented by 0 which has width $v = n > 8$. Hence, from (8.7) we see that the function $F^{(\kappa=0)}_{n|1}$ either has a simple pole at 0 or is a constant. Suppose the former were true, then $f_{n|1}$ defines a degree one function from $X_0(n)$ to $\mathbb{P}^1$, implying that $X_0(n)$ has genus zero as a Riemann surface. For $n \in \{11, 14, 15\}$ however, the genus of $X_0(n)$ is one, so we conclude that $f_{n|1}$ is constant in these cases.

☐

Lemma 8.4. *Suppose that $n \in \{10, 12\}$ and $h = 2$. Then $S_{n|h} = 0$.*

Proof. The only cusp of $\Gamma_0(20)$ that has width greater than 8 is that represented by 0, and similarly for $\Gamma_0(24)$, so in either case the only possible pole of $f_{n|h}$ is at the image under $\varphi_{nh} : \mathbb{H}^* \to X_0(nh)$ of the zero cusp. The degree of the pole, if the pole exists, is given by (8.6) with $2v = u = 2n$ and $m = 1$ equals 1, and so must be 1. In this case $f_{n|h} : X_0(nh) \to \mathbb{P}^1$ is a degree one map, implying that $X_0(nh)$ has genus zero. But this is not the case for $nh = 20$ or $nh = 24$. We conclude that $f_{n|h}$ has no poles and is thus constant. The claim follows.

☐

For the remaining three classes with $(n,h) = (23,1), (21,3), (12,12)$, the modular curves $X = X_0(nh)$ have higher genera and we will use the Riemann-Roch theorem:

$$\dim K_X(D) = \dim \Omega_X(D) + \deg(D) + 1 - \text{genus}(X) \tag{8.10}$$

to gain control over the allowed properties of the function $f_{n|h}$. Here D denotes a divisor on the Riemann surface X, $K_X(D)$ denotes the vector space of meromorphic functions f with the divisor (f) given by its zeros and poles satisfying $(f)+D \geq 0$, and $\Omega_X(D)$ denotes the space of holomorphic differentials ω on X satisfying $(\omega) - D \geq 0$.

Lemma 8.5. *If* $n = 23$ *then* $S_{n|1} = \lambda_n \eta^3$ *for some* $\lambda_n \in \mathbb{C}$.

Proof. The only cusp of $\Gamma_0(23)$ with width larger than 8 is the one with width 23 represented by 0, and hence the function $f_{23|1}$ can possibly have a pole only at $x \in X$ for x the image of 0 under φ_{23} (cf. (8.3)). More specifically, from (8.7) we see that $f_{23|1}$ is in the vector space $K_X(D)$ for $D = 2x$. Since the genus of $X_0(23)$ is 2, according to the Riemann-Roch formula (8.10) we have $\dim K_X(D) = 1 + \dim \Omega_X(D)$. We would like to prove the claim by showing that $\dim \Omega_X(D) = 0$ and hence $\dim K_X(D) = 1$, and therefore a function in $X(D)$ must be a constant function.

To show there is no holomorphic differential with a double zero at x, let us consider first the divisor $D' = x$. We know that $K_{D'}(X) \cong \mathbb{C}$, or else there would be a meromorphic function on X with a simple pole at x and no other poles, hence an isomorphism between X and $\mathbb{P}^1$, which is impossible given that the genus of X is 2. From this we conclude from the Riemann-Roch formula that $\dim \Omega_X(D') = \dim K_{D'}(X) = 1$ and there is a one-dimensional space of holomorphic differentials ω with $(\omega) - x \geq 0$. Now such a holomorphic differential is given by the weight 2 cusp form $\eta^2(\tau)\eta^2(23\tau)$, which is a holomorphic differential with a simple zero at x and a simple zero at the image of the infinite cusp under the canonical map $\mathbb{H}^* \to X_0(23)$ and no other zeros. From this we conclude $\dim \Omega_X(D) = 0$ and hence $\dim K_X(D) = 1$. This completes the proof. □

Lemma 8.6. *Suppose that* $n = 21$ *and* $h = 3$. *Then* $S_{n|h} = 0$.

Proof. For $n = 21$, by direct calculation one can check that the only cusp for which (8.8) does not hold is the cusp represented by 0. It corresponds to a cusp with width 63 of $\Gamma_0(63)$ and there are no other cusps of $\Gamma_0(63)$ at which $f_{21|3}$ can have a pole. Moreover, from (8.6) for the zero cusp, with

$$\tfrac{63}{21}\{\{\tfrac{21}{8} + \tfrac{1}{3}\}\} - \tfrac{63}{8} = -5$$

equation (8.6) shows that $f_{n|h}$ is in $K_D(X = X_0(63))$ with $D = 5x$, $x = \varphi_{63}(0)$.

The genus of $X_0(63)$ is 5 so according to the Riemann–Roch theorem (cf. (8.10)) we have $\dim K_X(D) = \dim \Omega_X(D) + 1$, where $\Omega_X(D)$ denotes the space of holomorphic differentials on X vanishing to order at least 5 at x. We claim that $\dim \Omega_X(D) = 0$. To show this, suppose $g \in S_2(\Gamma_0(63))$ corresponds to an element of Ω_X with a zero of order 5 at x. Since the Fricke involution w_{63} normalises $\Gamma_0(63)$, it induces a linear automorphism of $S_2(\Gamma_0(63))$. Then $g|_{1,2}\sigma$ is an element of $S_2(\Gamma_0(63))$ that is $\mathcal{O}(q^6)$ as $\tau \to i\infty$. By inspection there is no such cusp form of weight 2 for $\Gamma_0(63)$. We conclude that $\dim K_X(D) = 1$. Since $K_X(D)$ includes constant functions, our $f_{n|h}$ is constant, and the desired result follows.

□

Lemma 8.7. *Suppose that* $n = h = 12$. *Then* $S_{n|h} = 0$.

Proof. The group $\Gamma_0(nh) = \Gamma_0(144)$ has 9 cusps with widths exceeding 8: there are four with of width 9, having representatives 1/4, 3/4, 1/8 and 1/16, there are three of width 16, having representatives 1/3, 2/3 and 1/9, and there is one of width 36, with representative 1/2, and one of width 144, with representative 0. The four cusps of width 9 fuse under the action of $\Gamma_0(n) = \Gamma_0(12)$, as do the three cusps of width 16.

First we would like to see at which of these cusps of $\Gamma_0(144)$ the function $f_{n|h}$ could possibly have a pole. From the above data, we only need to look at the cusps of $\Gamma_0(12)$. From the formula (8.9) we see that $f_{n|h}$ is either constant or has a pole at the image of the zero cusp, as the pole-free condition (8.8) is satisfied at the all other cusps of $\Gamma_0(12)$. Moreover, putting $v = h = 12$, $m = 1$, $u = 144$ in the equation (8.6), we see that $f_{n|h}$ can only have a pole of order 11 if it is not a constant.

To show that a pole of order 11 cannot happen, first we will compute $\dim K_X(D) = 2$ for $X = X_0(144)$ and $D = 11x$ with $x = \varphi_{144}(0)$. The genus of $X_0(144)$ is 13 so according to the Riemann–Roch theorem (cf. (8.10)) we have $\dim K_X(D) = \dim \Omega_X(D) - 1$ where $\Omega_X(D)$ denotes the space of holomorphic differentials on X vanishing to order at least 11 at x. As in the proof of Lemma 8.6 we observe that elements of $\Omega_X(D)$ are in correspondence with cusp forms of weight 2 on $\Gamma_0(144)$ that are $\mathcal{O}(q^{12})$ as $\tau \to i\infty$. By inspection there is a three dimensional space of such cusp forms so $\dim K_X(D) = 2$.

Next set $D' = 10x$. We claim that then $K_X(D) = K_X(D')$. For certainly $K_X(D')$ is a subspace of $K_X(D)$ and their dimensions coincide since the Riemann–Roch theorem implies $\dim K_X(D') = \dim \Omega_X(D') - 2$ and $\dim \Omega_X(D') = 4$ by inspection of $S_2(\Gamma_0(144))$. We conclude that if $f_{n|h}$ is in $K_X(D)$ then it is also in $K_X(D')$ and then we must have $C_{n|h}(k = 0) = 0$ in (8.6), so that $f_{n|h}$ has no poles at any cusp and is therefore constant. This completes the proof. ☐

Remark. It is fascinating that the space $S_{3/2,\epsilon^3}(\Gamma_0(n))$ of cusp forms of weight 3/2 on $\Gamma_0(n)$ with multiplier system coinciding with that of η^3 is one-dimensional and spanned by η^3, whenever n is the order of an element of M_{23}. This is demonstrated by the above Lemmas taken together, while the same is certainly not true for a generic positive integer n.

For $n = 9$ and $n = 10$, for instance, it's possible that $F_{n|1} = S_{n|1}/\eta^3$ could define a degree one map $X_0(n) \to \mathbb{P}^1$, but this is no contradiction since $X_0(n)$ has genus zero for $n = 9$ and $n = 10$. Indeed, numerical approximations suggest that $S_{9|1}(\tau) = \eta(\tau)^3 + 3\eta(9\tau)^3$ and

$$S_{10|1}(\tau) = C\left(2\eta(\tau)^3 + 7\frac{\eta(2\tau)\eta(10\tau)^3}{\eta(5\tau)}\right)$$

for some $C \in \mathbb{C}$. Observe that the functions $\eta(\tau)^3/\eta(9\tau)^3$ and $\eta(\tau)^3\eta(5\tau)/\eta(2\tau)\eta(10\tau)^3$ are hauptmoduln for $\Gamma_0(9)$ and $\Gamma_0(10)$, respectively, and so $S_{3/2,\epsilon^3}(\Gamma_0(n))$ is not one dimensional for $n = 9$ or $n = 10$. Indeed, there are no M_{24} classes corresponding to $(n, h) = (9, 1)$ or $(10, 1)$.

Also in the case that $h > 1$, it is extremely non-trivial that for all the pairs (n, h) arising from M_{24}, just the correct combinations of genus, cusp widths and multipliers conspire to force the vanishing of $S_{n|h}$.

8.2 Determining R

First we will establish the identity $H_g = -2R_{n|1}$, for $n = n_g$ and for those g with $h_g = 1$. Recall from the discussion in §2 that these are the M_{24} classes whose action on the natural permutation representation has at least one fixed point (the M_{23} classes). To do this, we consider the function

$$G_{n|1} = \eta^3 \Big(\chi(g)\, R_{n|1} - \lambda_n H_g \Big) \tag{8.11}$$

with g being an M_{24} class with $n_g|h_g = n|1$. By construction, it has $G_{n|1} = (\chi(g) + 2\lambda_n) + \mathcal{O}(q)$ near the infinite cusp.

In §3 we saw that H_g is a weight 1/2 mock modular form for $\Gamma_0(n)$ with shadow $\chi(g)\eta^3$ and the same multiplier as η^{-3}. On the other hand, in §7 we have proven that the Rademacher sums $R_{n|1}$ are also weight 1/2 mock modular forms for $\Gamma_0(n)$ with the same multiplier. Moreover, it has as its shadow $S_{n|1}$, which was proven to be given by $\lambda_n \eta^3$ for some $\lambda_n \in \mathbb{C}$ in the previous subsection. Taken together, we see that the function $G_{n|1} : \mathbb{H}^* \to \mathbb{P}^1$ is a weight 2 modular form of $\Gamma_0(n)$ with trivial multiplier.

Using the properties of such weight 2 modular forms, we will be able to show that $G_{n|1} = 0$ for the values of n of interest to us. In particular, we have $\chi(g) + 2\lambda_n = 0$ as the constant term in the expansion near the infinite cusp. Since $\chi(g) \neq 0$ when $h_g = 1$, this implies $H_g = -2R_{n|1}$ for $g \in M_{24}$ with $h_g = 1$.

Lemma 8.8. *Let g be the identity element of M_{24} and set $n = n_g = 1$ and $h = h_g = 1$. Then $H = H_g = -2R_{n|h}$.*

Proof. From the above discussion we know the function

$$G_{n|1} = \eta^3 \left(24 R_{n|1} - \lambda_n H_g \right)$$

is a modular form of weight 2 for $\mathrm{SL}_2(\mathbb{Z})$ with trivial multiplier. There are no non-zero modular forms of weight 2 for $\mathrm{SL}_2(\mathbb{Z})$ so $G_{n|1}$ vanishes identically and by the above argument this proves the lemma. ☐

To prove $G_{n|1} = 0$ for $n > 1$ we need to study its behaviour near the other cusp representatives of $\Gamma_0(n)$. For a cusp $\kappa = \sigma\infty$, $\sigma \in \mathrm{SL}_2(\mathbb{Z})$, with width v in $\Gamma_0(n)$, we shall again consider its transformation under a scaling element $\tilde{\sigma}$ (cf. (7.3))

$$G^{(\kappa)}_{n|1}(\tau) = G_{n|1}\big|_{1,2} \sigma \left(v(\tau + \beta) \right) .$$

To be more precise, we would like to know the form of the Fourier expansion of $G^{(\kappa)}_{n|1}(\tau)$. As will be shown explicitly in Appendix E, we have

$$G^{(\kappa)}_{n|1}(\tau) = \mathcal{O}(q^{v/8+\Delta})$$

as $\tau \to i\infty$ with $\Delta \geq 0$, which corresponds to the fact that both $R_{n|1}$ and H_g have no pole at any cusp other than the infinite cusp. By construction, $G^{(\kappa)}_{n|1}$ is invariant under $\tau \to \tau + 1$ and hence the above argument shows

$$G^{(\kappa)}_{n|1}(\tau) = \mathcal{O}(q^{\lceil v/8 \rceil}) , \tag{8.12}$$

where the ceiling function $\lceil x \rceil$ gives the smallest integer not less than $x \in \mathbb{R}$.

Lemma 8.9. *Suppose that $g \in M_{24}$ has prime order $n = n_g$ and $h = h_g = 1$. Then $H_g = -2R_{n|h}$.*

Proof. The primes dividing the order of M_{24} are exactly those primes p for which $(p+1)|24$ so $n \in \{2,3,5,7,11,23\}$. For these n, the only non-infinite cusp of $\Gamma_0(n)$ is the one represented by 0, with width n. Take the Fricke involution w_n to be the scaling element. From (8.12) we conclude that $G^{(\kappa=0)}_{n|1}$ is $\mathcal{O}(q)$ for $n \in \{2,3,5,7\}$ and $\mathcal{O}(q^2)$ for $n = 11$ and $\mathcal{O}(q^3)$ for $n = 23$.

Since the Fricke involution w_n for $\Gamma_0(n)$ normalises $\Gamma_0(n)$, $G^{(\kappa=0)}_{n|1}$ also belongs to $M_2(\Gamma_0(n))$. The space $M_2(\Gamma_0(n))$ of such weight 2 modular forms contains the space $S_2(\Gamma_0(n))$ of weight 2 cusp forms. The dimension of the latter is given by the genus of the modular curve $X_0(n)$. A complement in the former to the latter is spanned by $\{\Lambda_m,\ m>1,\ m|n\}$ with

$$\begin{aligned}\Lambda_m(\tau) = m\,q\frac{d}{dq}\log(\eta(n\tau)/\eta(\tau)) &= \frac{m}{24}\Big(m\,E_2(m\tau) - E_2(\tau)\Big)\\ &= \frac{m(m-1)}{24} + \sum_{k=1}^{\infty} m\sigma(k)(mq^{mk} - q^k)\end{aligned} \tag{8.13}$$

where $\sigma(k)$ is the sum of the divisors of k and E_2 is the holomorphic quasi-modular Eisenstein series of weight 2.

When n is prime, obviously the complement in $M_2(\Gamma_0(n))$ to $S_2(\Gamma_0(n))$ is one-dimensional and spanned by Λ_n. Since the genus of $X_0(n)$ is zero for $n \in \{2,3,5,7\}$, the space $M_2(\Gamma_0(n))$ is one-dimensional and spanned by Λ_n. Now, with Λ_n having non-vanishing constant term, the statement $G^{(\kappa=0)}_{n|1}(\tau) = \mathcal{O}(q)$ implies that $G^{(\kappa=0)}_{n|1}(\tau)$ vanishes identically.

If $n = 11$ then $M_2(\Gamma_0(n))$ is two dimensional, spanned by Λ_n and the cusp form $\eta_g(\tau) = \eta(\tau)^2\eta(11\tau)^2$, and since Λ_n and η_g have non-vanishing coefficients of q^0 and q^1, respectively, we conclude that $G^{(\kappa=0)}_{n|1}(\tau) = \mathcal{O}(q^2)$ implies the vanishing of $G^{(\kappa=0)}_{n|1}$ in this case also. When $n = 23$, the space $M_2(\Gamma_0(n))$ is spanned by Λ_n, a cusp form $\phi_{23,1} = \eta_g(\tau)^2 = \eta(\tau)^2\eta(23\tau)^2$ and a further cusp form

$$\phi_{23,2}(\tau) = \frac{\eta(\tau)^3\eta(23\tau)^3}{\eta(2\tau)\eta(46\tau)} + 4\eta(\tau)\eta(2\tau)\eta(23\tau)\eta(46\tau) + 4\eta(2\tau)^2\eta(46\tau)^2. \tag{8.14}$$

Inspecting the Fourier expansions of Λ_n, $\phi_{23,1}$ and $\phi_{23,2}$ we see that there is no non-zero linear combination that is $\mathcal{O}(q^3)$. We conclude that $G^{(\kappa=0)}_{n|1}(\tau)$ vanishes identically for all the values of n in question. So $G_{n|1}$ also vanishes and the desired result follows. ☐

Lemma 8.10. *Suppose that $g \in M_{24}$ has order $n = n_g$ a product of two distinct primes and $h = h_g = 1$. Then $H_g = -2R_{n|h}$.*

Proof. According to Table 1 the values of n in question are 6, 14 and 15. From the discussion before (8.13) we see that the complement to $S_2(\Gamma_0(n))$ in $M_2(\Gamma_0(n))$ is three-dimensional and spanned by the weight two modular forms $\Lambda_{e_1}, \Lambda_{e_2}$ and Λ_{e_3}, where $e_1 < e_2 < e_3 = n$ are the three (exact) divisors of n that are larger than 1. Therefore, we have

$$G_{n|1} = \alpha_1\Lambda_{e_1} + \alpha_2\Lambda_{e_2} + \alpha_3\Lambda_{e_3} + \phi$$

with $\phi \in S_2(\Gamma_0(n))$. In other words, ϕ vanishes when $n = 6$ and is a multiple of η_g when $n = 14, 15$, in accordance with the fact that the genera of $X_0(6)$, $X_0(14)$, $X_0(15)$ are 0, 1, 1, respectively.

For each e_i there is a corresponding cusp of $\Gamma_0(n)$ with width e_i, which is represented by $\kappa_i = \frac{e_i}{n}$. These are the only cusps of $\Gamma_0(n)$ apart from the infinite cusp, as listed in Table 3. As discussed before (cf. (8.12)), the corresponding function $G^{(\kappa_i)}_{n|1}(\tau)$ has the asymptotic behaviour

$$G^{(\kappa_i)}_{n|1}(\tau) = \mathcal{O}(q^{\Delta+\frac{e_i}{8}}) \ .$$

Consider a scaling element which is an Atkin-Lehner involution w_{e_i} with $\kappa_i = w_{e_i}\infty$ (note that it is different from the Fricke involution when $e_i \neq n$). With help from the identity

$$\tfrac{1}{e}\Lambda_e|_{1,2}w_f = \tfrac{1}{e*f}\Lambda_{e*f} - \tfrac{1}{f}\Lambda_f \tag{8.15}$$

where $e*f = ef/(e,f)^2$, we obtain three equations on α_i from the vanishing of the constant terms in $G^{(\kappa_i)}_{n|1}(\tau)$. Solving the resulting linear system we quickly deduce that $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and hence $G_{n|1}(\tau)$ is given by the cusp form ϕ of $\Gamma_0(n)$.

Since $X_0(6)$ has genus zero we conclude that $G_{n|1}(\tau) = 0$ for $n = 6$. For $n = 14, 15$, to prove that $G_{n|1}(\tau) = \phi = 0$ let us focus on the 0 cusp with width n. Take the scaling element to be the Fricke involution w_n. From the fact that the cusp form η_g is an eigenfunction (with eigenvalue -1) of w_n we get $G^{(\kappa)}_{n|1}(\tau) = C\eta_g = -Cq + \mathcal{O}(q^2)$ for some $C \in \mathbb{C}$. But from (8.12) we see that $G^{(\kappa)}_{n|1}(\tau) = \mathcal{O}(q^2)$ for $v = n = 14, 15$. Hence we conclude $G_{n|1} = C = 0$. □

Lemma 8.11. *Suppose that $g \in M_{24}$ is such that $h = h_g = 1$ and $n = n_g$ is 4 or 8. Then $H_g = -2R_{n|h}$.*

Proof. If $n = 4$ then $\dim M_2(\Gamma_0(n)) = 2$ and $M_2(\Gamma_0(n))$ is spanned by the modular forms Λ_2 and Λ_4 (cf. (8.13)). If $n = 8$ then $\dim M_2(\Gamma_0(n)) = 3$ and $M_2(\Gamma_0(n))$ is spanned by Λ_2, Λ_4 and Λ_8. In both cases $\dim S_2(\Gamma_0(n)) = 0$.

Consider the case that $n = 4$, where we have $G_{n|1} = \alpha\Lambda_2 + \beta\Lambda_4$. There are two non-infinite cusps: one represented by 0 with width 4 and the other represented by 1/2 with width 1. Taking as their scaling elements w_4 and $T^{1/2}w_4$ respectively, we get two equations on the coefficients α, β from the requirement that both $G^{(\kappa=0)}_{n|1}$ and $G^{(\kappa=1/2)}_{n|1}$ are $\mathcal{O}(q)$, which force both α and β and hence $G_{n|1}$ to vanish. In arriving at these equations, we have used the identity (8.15) and

$$E_2(\tau + 1/2) = -E_2(\tau) + 6E_2(2\tau) - 4E_2(4\tau) \ .$$

The argument for $n = 8$ is very similar. We have $F = \alpha\Lambda_2 + \beta\Lambda_4 + \gamma\Lambda_8$ for some $\alpha, \beta, \gamma \in \mathbb{C}$. We may take 0, 1/2 and 1/4 as representatives for the three non-infinite cusps of $\Gamma_0(8)$. We may take w_8, $T^{1/2}w_8$ and $w_8T^{1/2}w_8$, respectively, as scaling elements for these cusp representatives. Using the identities given above we compute expressions for the constant term at each non-infinite cusp as linear equations in α, β and γ. From the vanishing of each constant term we deduce that $\alpha = \beta = \gamma = 0$ and hence $G_{n|1} = 0$ also. Then the required identity $H_g = -2R_{n|h}$ follows as before. □

The above Lemmas 8.8-8.11, when taken together, show that $H_g = -2R_{n_g|1}$ for $g \in M_{24}$ with $h_g = 1$ Now we will continue to show the identity for the remaining M_{24} classes with $n_g = n$, $h_g = h > 1$.

Lemma 8.12. *Let g be an element of M_{24} that has $h_g > 1$ and $n_g \neq 21$. Then $H_g = -2R_{n|h}$.*

Proof. To prove the identity $-2R_{n|h} = H_g$ for the remaining classes $[g]$ with $h_g > 1$, we consider the following function

$$K_{n|h} = \frac{R_{n|h}}{H_g} \ .$$

Since both $R_{n|h}$ and $H_g = -\tilde{T}_g/\eta^3$ are known to be weight 1/2 modular forms which transform with a multiplier $\rho_{n|h}$ on $\Gamma_0(n)$, we conclude that $K_{n|h}$ is a modular function on $\Gamma_0(n)$. As in the previous subsection, such a modular function has to be either constant or have a pole. For $g \in M_{24}$ with $h_g > 1$ and $n_g \neq 21$, the McKay–Thompson series H_g are all given by certain η-quotients (cf. Table 2). Since η does not have a zero in $\mathbb{H}$, we only have to check that $K_{n|h}$ has no pole at any cusp of $\Gamma_0(n)$. At the infinite cusp we have $R_{n|h} = q^{-1/8} + \mathcal{O}(q^{7/8})$ and $H_g = -2q^{-1/8} + \mathcal{O}(q^{7/8})$ and hence $K_{n|h}$ has no pole at the infinite cusp.

To see that it is also bounded near a non-infinite cusp $\kappa = \sigma\infty = \tilde{\sigma}\infty$, $\sigma \in \mathrm{SL}_2(\mathbb{Z})$, with a scaling element $\tilde{\sigma}$, we consider the function

$$G^{(\kappa)}_{n|h} = G^{(\kappa)}_{n|h}\big(\sigma\left(v(\tau+\beta)\right)\big) = \frac{R^{(\kappa)}_{n|h}(\tau)}{\tilde{T}_g/\eta^3|_{\epsilon^{-3},1/2}\,\sigma(v(\tau+\beta))}$$

where $R^{(\kappa)}_{n|h}(\tau)$ has been shown in Appendix E to have an expansion

$$R^{(\kappa)}_{n|h}(\tau) = R^{(\kappa)}_{n|h}\big|_{\epsilon^3,1/2}\sigma\left(v(\tau+\beta)\right) = \sum_{k=1}^{\infty} c^{(\kappa)}_{\Gamma,\rho}(k-\nu)\,\mathrm{e}(\tau(k-\nu)) \ , \tag{8.16}$$

where $\nu = \{\{\frac{m}{h} + \frac{v}{8}\}\}$ as explained in (7.5). With direct calculation, one can check that also

$$\frac{\tilde{T}_g}{\eta^3}\big|_{\epsilon^{-3},1/2}\,\sigma(v(\tau+\beta)) = a_g\, q^{1-\nu} + \mathcal{O}(q^{2-\nu})$$

at every non-infinite cusp κ of $\Gamma_0(n)$ with some $a_g \in \mathbb{C}$, $a_g \neq 0$.

This shows that $K_{n|h}$ is a modular function with no pole and hence constant, which is $-1/2$ from its value at the infinite cusp.

□

Lemma 8.13. *Suppose that $g \in M_{24}$ is such that $n = n_g = 21$ and $h = h_g = 3$. Then $H_g = -2R_{n|h}$.*

Proof. Set $\phi_{21,1}(\tau) = \eta(7\tau)^3/\eta(3\tau)\eta(21\tau)$ and $\phi_{21,2}(\tau) = \eta(\tau)^3/\eta(3\tau)^2$. Then both $\phi_{21,1}$ and $\phi_{21,2}$ are (meromorphic) modular forms of weight 1/2 for $\Gamma_0(21)$ with multiplier system $\rho\epsilon^{-3}$ where $\rho = \rho_{21|3}$, so that

$$G = \frac{AR_{n|h} + B\phi_{21,1}}{\phi_{21,2}} \quad , \quad A, B \in \mathbb{C}$$

is a modular function for $\Gamma_0(21)$ whose only poles are at the non-infinite cusps of $\Gamma_0(21)$. Since $R_{n|h}, \phi_{21,2}, \phi_{21,2}$ all have the expansion $q^{-1/8} + \mathcal{O}(q^{7/8})$ near the infinite cusp, we see that G has no pole at the infinite cusp. For $e_{1,2,3}$ with $e_1 < e_2 < e_3 = 21$ the exact divisors of 21 that are larger than 1 let w_{e_i} be an element of the Atkin–Lehner involution of $\Gamma_0(21)$ associated to e_i. Then the elements w_{e_i} furnish scaling elements at the respective representatives $\kappa_i = w_{e_i}\infty$ for the three non-infinite cusps of $\Gamma_0(21)$, represented by $\frac{e_i}{n}$ having width e_i. Using the fact that the slash operator $|_{1/2,\epsilon} w_{e_i}$ maps $\eta(e_j\tau)$ to $\eta(e_i * e_j\tau)$ (up to a non-vanishing scale factor), we see that the function $G^{(\kappa_i)} = G(w_{e_i}\tau)$ remains bounded as $\tau \to i\infty$ except possibly in the case that $e_3 = 21$, in which case we have

$$G^{(\kappa=0)}(\tau) = G(w_{21}\tau) = \left(A' R^{(\kappa=0)}_{n|h}(\tau) + B' \frac{\eta(3\tau)^3}{\eta(7\tau)\eta(\tau)} \right) \frac{\eta(7\tau)^2}{\eta(21\tau)^3} \,,\ A', B' \in \mathbb{C}.$$

From (8.16) we see that $G^{(\kappa=0)}(\tau)$ has a Fourier expansion of the form $Cq^{-2} + \mathcal{O}(q^{-1})$. Setting $D = 2x$ where x is the image of $w_{21}\infty = 0$ under the natural map $\varphi_{21} : \mathbb{H}^* \to X = X_0(21)$ we see that f must belong to the space $K_X(D)$ of meromorphic functions on $X = X_0(21)$ having a pole of order at most 2 at x and no other poles. The genus of $X_0(21)$ is 1 so according to the Riemann–Roch theorem (cf. (8.10)) we have $\dim K_X(D) = \dim \Omega_X(D) + 2$ where $\Omega_X(D)$ denotes the space of holomorphic differentials on X vanishing to order at least 2 at x. The full space Ω_X of holomorphic differentials on X is in correspondence with $S_2(\Gamma_0(21))$, and this one-dimensional space is in turn spanned by the L-series for the elliptic curve $y^2 + xy = x^3 + x$. This cusp form has a Fourier expansion of the form $q + \mathcal{O}(q^2)$ and is an eigenform for w_{21} so we conclude that it does not belong to $\Omega_X(D)$, and thus $\dim \Omega_X(D) = 0$ and $\dim K_X(D) = 2$. Now $K_X(D)$ includes constant functions on X and it also includes the non-constant function $\phi_{21,1}/\phi_{21,2}$, so it is spanned by these functions, and we may conclude that $R_{n|h} = C_1\phi_{21,1} + C_2\phi_{21,2}$ for some $C_1, C_2 \in \mathbb{C}$. By comparison of polar terms in the Fourier expansions of $R_{n|h}$, $\phi_{21,1}$ and $\phi_{21,2}$ we have $C_1 + C_2 = 1$. Now consider the expansion of $R_{n|h}$ at the cusp represented by $w_7\infty$ (which is also represented by 1/3). We have $(\phi_{21,1}|w_7)(\tau) = D_1 q^{-21/24} + \mathcal{O}(q^{3/24})$ and $(\phi_{21,2}|w_7)(\tau) = D_2 q^{-21/24} + \mathcal{O}(q^{3/24})$ and the fact that $R_{n|h}$ has no poles away from the infinite cusp implies that $C_1 + 7C_2 = 0$. We conclude that $6R_{n|h} = 7\phi_{21,1} - \phi_{21,2}$. From Table 2 we have the explicit expression $3H_g = \phi_{21,2} - 7\phi_{21,1}$. This proves the lemma. □

9 Spectral Theory

In this section we would like to demonstrate the convergence of the Selberg–Kloosterman zeta function $Z_{0,k}(s, \rho\epsilon^{-3})$, defined in (5.6) and (5.7), of $\Gamma = \Gamma_0(n)$ for n a positive integer. As before, we take $\psi = \epsilon^{-3}\rho_{n|h}$ for some h dividing both 24 and n (cf. (2.5)). Combined with the discussions in §5 and §6, Theorem 9.1 completes the proof of the convergence of the Rademacher sum (4.8).

Let w be a positive real number and let ψ be a multiplier system of weight w for Γ. We say a function $g(z)$ on $\mathbb{H}$ is *automorphic for* Γ *of weight* w *with respect to* ψ if

$$\psi(\gamma) g(\gamma z) \exp(iw \arg(\operatorname{jac}(\gamma, z))/2) = g(z)$$

for all $\gamma \in \Gamma$, where $\arg(\cdot)$ is defined so that $-\pi < \arg(z) \leq \pi$ for $z \neq 0$, and write $\mathcal{H}_{\psi,w}$ for the Hilbert space consisting of those automorphic functions which also satisfy the growth condition

$$\iint_{\Gamma\backslash\mathbb{H}} |g(\tau)|^2 \frac{\mathrm{d}x\mathrm{d}y}{y^2} < \infty$$

where $x = \Re(\tau)$ and $y = \Im(\tau)$ for $z \in \mathbb{H}$. We consider the operator Δ_w on $\mathcal{H}_{\psi,w}$ given by

$$\Delta_w = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - iwy \frac{\partial}{\partial x}.$$

According to [33] if ψ is non-trivial on every parabolic subgroup of Γ then Δ_w has a discrete spectrum of eigenfunctions in $\mathcal{H}_{\psi,w}$ and there exists a complete orthonormal system $\{u_j(z)\}$ of eigenfunctions satisfying

$$\Delta_w u_j(z) + \lambda_j u_j(z) = 0$$

for a set $\{\lambda_j\}$ of eigenvalues which are all real. (Roelcke proves in [34] that Δ_w has a unique self-adjoint extension to $\mathcal{H}_{\psi,w}$.) Further, the λ_j are all positive except possibly for a finite number which are of the form

$$(w/2 - l)(1 + l - w/2) \tag{9.1}$$

where l is a non-negative integer less than $w/2$. If there is an eigenvalue of the form (9.1) then, negative or otherwise, it's multiplicity is equal to the dimension of the space $S_{w-2l}(\Gamma, \psi)$ of (holomorphic) cusp forms for Γ of weight $w - 2l$ with respect to ψ.

By the trivial estimate $|S(m, k, c, \psi)| < c$ the series (5.7) defining the Selberg–Kloosterman zeta function $Z_{m,k}(s, \psi)$ converges absolutely and locally uniformly for $\Re(s) > 1$ and thus defines a holomorphic function in this domain. Selberg demonstrates in [33] that this function admits an analytic continuation to a function meromorphic in the entire s-plane that is holomorphic for $\Re(s) > 1/2$ except possibly for a finite number of simple poles lying on the real segment $1/2 < s < 1$. Poles on this segment can only occur at points of the form $1/2 + \sqrt{1/4 - \lambda_j}$ where $0 < \lambda_j < 1/4$ and λ_j is an eigenvalue of Δ_w, and such a pole does occur if λ_j is not of the form (9.1) and the coefficients of $\mathrm{e}((m - \alpha)x)$ and $\mathrm{e}((n - \alpha)x)$ in the Fourier expansion (with respect to x) of the corresponding eigenfunction $u_j(z)$ are both non-zero. If λ_j is of the form (9.1) with $l = \lfloor w/2 \rfloor$ then there is a pole at $s = 1 + \lfloor w/2 \rfloor - k/2$ just in the case that there is a cusp form $f(\tau) \in S_{w-2\lfloor w/2 \rfloor}(\Gamma, \psi)$ for which the coefficients of $q^{m-\alpha}$ and $q^{n-\alpha}$ (in its Fourier expansion with respect to τ) are both non-zero.

For the case of relevance for us we have $w = 1/2$ and so all the λ_j will be positive. Roelcke proves in [34] that the minimal value of a λ_j is $3/16$ in case $w = 1/2$ (see [35] for a nice exposition) so we conclude that there are no poles for $Z_{0,k}(s, \rho\epsilon^{-3})$ in $\Re(s) > 3/4$. The only possible value for l in (9.1) is $l = 0$ so that the occurrence of a pole at $s = 1 + \lfloor w/2 \rfloor - w/2 = 3/4$ depends upon the existence of a cusp form $f \in S_w(\Gamma, \psi)$ for which the coefficients of $q^{0-\alpha}$ and $q^{k-\alpha}$ are both non-zero. We have $-\alpha = -1/8$ so there is no such cusp form and we conclude that $Z_{0,k}(s, \rho\epsilon^{-3})$ is in fact analytic in $\Re(s) > 3/4 - \delta$ for some $\delta > 0$ and has finitely many simple poles lying on the line segment $1/2 < s < 3/4$.

It remains to show that the value of the analytically continued function $Z_{0,k}(s, \rho\epsilon^{-3})$ at $s = 3/4$ is indeed the sum

$$\sum_{c>0} \frac{S(0, k, c, \rho\epsilon^{-3})}{c^{3/2}}$$

that we expect it is. To establish this we adopt the approach presented in §2 of [36][3] which is an adaptation of a technique commonly used in proving the prime number theorem (cf. [38]). (For a very crisp application of what is essentially the same argument, see the proof of Theorem 2 in [32].) Fix a group Γ, a multiplier $\psi = \rho\epsilon^{-3}$, and an integer k. We will write $Z(s) = Z_{0,k}(s, \rho\epsilon^{-3})$ to ease notation. Define

$$\Sigma(x) = \sum_{0<c<x} \frac{S(0, k, c, \rho\epsilon^{-3})}{c^{3/2}}$$

---

[3] If you consult this reference please see also §II of [37] for some important corrections.

for $x > 0$. We require to show that $\lim_{x\to\infty} \Sigma(x) = Z(3/4)$. Observe that for $a > 0$ and $x \neq c$ the integral

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \left(\frac{x}{c}\right)^t \frac{\mathrm{d}t}{t}$$

is 1 or 0 according as $x > c$ or $x < c$. Consequently we may rewrite $\Sigma(x)$ as

$$\Sigma(x) = \sum_{c>0} \frac{S(0,k,c,\rho\epsilon^{-3})}{c^{3/2}} \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \left(\frac{x}{c}\right)^t \frac{\mathrm{d}t}{t}$$

for non-integral values of x. Choose $\varepsilon > 0$ and set $a = 1/2 + \varepsilon$ so that $Z(t/2 + 3/4)$ converges absolutely for $\Re(t) = a$. For $T > 0$ define $\Sigma_T(x)$ by setting

$$\Sigma_T(x) = \frac{1}{2\pi i}\int_{a-iT}^{a+iT} Z(t/2+3/4)x^t \frac{\mathrm{d}t}{t}.$$

Then for the difference $\Sigma(x) - \Sigma_T(x)$ we have the upper bound

$$|\Sigma(x) - \Sigma_T(x)| < x^a \sum_{c>0} c^{-1-\varepsilon} \min\left(1, \frac{1}{T|\log(x/c)|}\right) \tag{9.2}$$

thanks to the following estimate, which may be found in §17 of [38].

$$\frac{1}{2\pi}\left|\int_{a-i\infty}^{a+i\infty} \left(\frac{x}{c}\right)^t \frac{\mathrm{d}t}{t} - \int_{a-iT}^{a+iT} \left(\frac{x}{c}\right)^t \frac{\mathrm{d}t}{t}\right| < \left(\frac{x}{c}\right)^a \min\left(1, \frac{1}{T|\log(x/c)|}\right)$$

Following exactly the argument of the last paragraph of p.171 of [36] we can replace the summation in (9.2) with the more explicit bound

$$|\Sigma(x) - \Sigma_T(x)| < K\frac{x^{1/2}}{T}(x^\varepsilon + \log x) \tag{9.3}$$

where K is some constant (that does not depend on our choice of n).

Consider now the integral

$$\frac{1}{2\pi i}\int_C Z(t/2+3/4)\, x^t\, \frac{\mathrm{d}t}{t}$$

where $C = C_0 + C_1 + C_2 + C_3$ is the positively oriented rectangle with corners $(\pm 1/2) + \varepsilon \pm iT$. For concreteness let us say that C_0 is the right most vertical component and C_1 is the upper most horizontal component, &c. Then the contribution of C_0 is exactly $\Sigma_T(x)$. We denote the contribution of C_i by I_i for $i \neq 0$. By the residue theorem we have

$$\Sigma_T(x) + I_1 + I_2 + I_3 = Z(3/4) + \sum_{j\in J} \varrho_j x^{2s_j - 3/2} \tag{9.4}$$

where $\{s_j\}_{j\in J}$ is an enumeration of the (finite) set of poles of $Z(s)$ in the interval $1/2+\varepsilon/2 < s < 3/4$ (we may choose $\varepsilon > 0$ so that $s = 1/2 + \varepsilon/2$ is not a pole) and ϱ_j is the residue of $Z(s)$ at $s = s_j$ divided by $2s_j - 3/2$.

We wish to determine what happens when $x \to \infty$ in (9.4). Observe, for a start, that the sum over j vanishes in this limit since $2s_j < 3/2$ for all $j \in J$. We claim that the integrals I_i also vanish in this limit. To estimate the I_i we utilise the growth condition

$$|Z(s)| = O\left(\frac{|s|^{1/2}}{\Re(s) - 1/2}\right)$$

as $|\Im(s)| \to \infty$, valid for $\Re(s) > 1/2$. Such a result was first established by Goldfeld–Sarnak in [32]. We have applied the generalisation found in [39] due to Pribitkin. Since $Z(s)$ converges absolutely for $\Re(s) > 1$ we have that $|Z(t/2 + 3/4)| = \mathcal{O}(1)$ as $|\Im(t)| \to \infty$ when $\Re(t) = 1/2 + \varepsilon$, and thus an application of the Phragmén–Lindelöf theorem (cf. [40], Thm. 14) yields

$$|Z(t/2 + 3/4)| < K_\varepsilon \, |\Im(t)|^{1/4 - \Re(t)/2 + \varepsilon/2} \tag{9.5}$$

for all t such that $-1/2 + \varepsilon \leq \Re(t) \leq 1/2 + \varepsilon$ and $|\Im(t)|$ is sufficiently large, for some constant K_ε depending only on Γ, $\psi = \rho\epsilon^{-3}$, k and ε. Applying (9.5) to the integrals I_i we obtain that $I_2 = \mathcal{O}(T^{1/2}x^{-1/2+\varepsilon})$, and both I_1 and I_3 are $\mathcal{O}(T^{-1}x^{1/2+\varepsilon})$ so long as $T < x^2$. Setting $T = x^{2/3}$, letting $x \to \infty$ in (9.4) and applying (9.3) gives us the desired result: $\lim_{x\to\infty} \Sigma(x) = Z(3/4)$. In particular, we have the following theorem:

Theorem 9.1. *Let $\Gamma = \Gamma_0(n)$ for some positive integer n, let h be a divisor of* $\gcd(n, 24)$ *and set $\rho = \rho_{n|h}$. Then for any positive integer n the Selberg–Kloosterman zeta function $Z_{0,k}(s, \rho\epsilon^{-3})$ converges at $s = 3/4$.*

10 Conclusion and Discussion

To summarise, motivated by the AdS/CFT correspondence in physics, we propose to replace the genus zero property by the more general property of Rademacher summability as the organising principle of the modular properties of moonshine phenomena. In monstrous moonshine, where the McKay–Thompson series are modular functions, the latter coincides with the former and is hence automatically satisfied, as shown in [10]. The two conditions decouple in the case of the recently conjectured relation between mock modular forms and M_{24}: on the one hand, it is neither necessary nor sufficient that a group Γ have genus zero for it to arise as Γ_g for some g in M_{24}, and on the other hand, we verify the Rademacher summability property in Theorem 4.1. In view of the above, the Rademacher summability property may replace the genus zero property, and applies to both monstrous moonshine and M_{24}. This also suggests that AdS/CFT considerations will lead to further elucidation of the moonshine phenomena, and we expect the powerful Rademacher machinery to prove useful in further study of the connection between modular objects and finite groups in general.

This result raises many interesting physical questions. In particular, what do we know about the gravity duals, if they exist, of 2d CFTs with sporadic symmetries? Let us first concentrate on the sigma models with target spaces involving $K3$ surfaces, as they are relevant for the structures studied in the present paper. Recall that one of the best known examples of the AdS/CFT correspondence conjectures that the two-dimensional CFT describing the Higgs branch of the system of Q_1 D1-branes bound to Q_5 D5-branes wrapping a $K3$ surface is dual to the type IIB string theory in the background $AdS_3 \times S^3 \times K3(Q)$ [12]. Here the radius of curvature of AdS_3 and S^3 is given by $(Q_1Q_5)^{1/4}$ in the six dimensional unit, and $K3(Q)$ is the "attractor K3" whose moduli are partially fixed by the requirement of minimising the BPS mass of the given charges, or equivalently, by the so-called attractor mechanism of supergravity. Note that not all $K3$ surfaces can be such an attractor $K3$ for a given set of charges. On the conformal field theory side, the Higgs branch is described by a 2d $\mathcal{N} = (4,4)$ SCFT with central charge $c = 6k$, $k = Q_1Q_5 + 1$, which is a deformation of the sigma model on the symmetric product S^kK3, as can be most easily understood by going to the D0-D4 duality frame [41].

This AdS/CFT dictionary makes it clear that a "semi-classical-like" AdS gravity description is only guaranteed to exist when we take the central charge $c = 6k$, or equivalently the AdS radius, to be very large. For this reason, it is in fact surprising that the Rademacher machinery works so well when applied to the partition functions for the $K3$ sigma model with $k = 1$, as demonstrated in the present paper. Two comments on this issue are in order here. First, analogous Rademacher formulas also exist for the symmetric product sigma model with higher k and we plan to report on this in more detail in the future [27]. In other words, the Rademacher machinery also works when we have very strong reasons to expect them to work. Second, recently there has been important progress in developing the localisation techniques for computing gravity path integrals (see, for example, [42]), and we have seen that in some cases the highly quantum path integral actually localises into the same form as if coming from a weakly coupled, semi-classical gravitational theory. For instance, very recently the Rademacher sum for the modular form $1/\eta(\tau)^{24}$, which is the partition function of a CFT with central charge as small as $c = 24$, has been reproduced from a localised gravity path integral [43]. This is very surprising since the corresponding black holes have zero macroscopic entropy (that is, the Bekenstein-Hawking entropy grows only linearly with the charges) and hence a priori we have no reason to expect the gravity path integral to take this "semi-classical" form. These results show that the Rademacher formulas are not really the monopoly of a large radius, semi-classical gravity description, although the former is most easily motivated in the semi-classical limit as we reviewed in §1, and suggests that it is not unlikely that a localisation calculation could similarly explain the Rademacher sum formulas derived in the present paper from a gravity point of view. In the case of monstrous moonshine, on the other hand, the question of whether a semi-classical theory of gravity dual to the moonshine module conformal field theories, or possibly their higher central charge cousins, is an interesting open puzzle [44].

We end with a few more comments concerning the Rademacher formulas studied in this paper. Firstly, it is remarkable that the Rademacher sums with such a simple polar term $-2q^{-1/8}$ compute not only the M_{24}-modules underlying the massive states in the $K3$ CFT but also the twisted Euler characteristics of $K3$ surfaces. The latter statement comes from the fact that the Rademacher machinery automatically computes the shadows of the resulting mock modular forms, which are in this case given by the twisted $K3$ Euler characteristics multiplied by the modular form $\eta(\tau)^3$. Finally, we would like to mention that none of the twisted elliptic genera which arise from the extra, non-M_{24} symmetries discussed in [45] lead to (mock) modular forms admitting analogous Rademacher sum expressions. Another related but logically independent observation is the following. If we were to twist the partition function counting the 1/4-BPS index of type II theory compactified on $K3 \times T^2$ with one of these non-M_{24} symmetries, the resulting candidate 4d twisted partition function would have a system of poles and zeros that is more intricate and complicated than the ones coming from twisting with M_{24} symmetries. This adds to the distinction between the M_{24} symmetries and the extra non-M_{24} discrete symmetries which are present in, for instance, the Gepner models, whose role in moonshine is still obscure.

Acknowledgements

We thank Atish Dabholkar, Rajesh Gopakumar, Tom Hartman, Özlem Imamoglu, Jan Manschot, Shiraz Minwalla, Ashoke Sen, Edward Witten, and Don Zagier for many useful discussions, and we thank Amanda Folsom and Wladimir de Azevedo Pribitkin for communication related to convergence. We also thank Alejandra Castro, Thomas Creutzig, Igor Frenkel, Christoph Keller, Alex Maloney, Andy Strominger and S.T. Yau for earlier discussions on related topics. M.C. would like to thank the Institute for Advanced Study for hospitality during part of this work, and the organisers and participants of "the School and Conference on Modular Forms and mock Modular Forms and their Applications in Arithmetic, Geometry and Physics" for stimulating discussions. The research of M.C. is supported in part by NSF grant DMS-0854971.

A Dedekind Eta Function

The *Dedekind eta function*, denoted $\eta(\tau)$, is a holomorphic function on the upper half-plane defined by the infinite product

$$\eta(\tau) = q^{1/24} \prod_{n\geq 1}(1-q^n)$$

where $q = \mathrm{e}(\tau) = e^{2\pi i\tau}$. It is a modular form of weight $1/2$ for the modular group $\mathrm{SL}_2(\mathbb{Z})$ with multiplier $\epsilon : \mathrm{SL}_2(\mathbb{Z}) \to \mathbb{C}^*$, which means that

$$\epsilon(\gamma)\eta(\gamma\tau)\,\mathrm{jac}(\gamma,\tau)^{1/4} = \eta(\tau)$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, where $\mathrm{jac}(\gamma,\tau) = (c\tau+d)^{-2}$. The *multiplier system* ϵ may be described explicitly as

$$\epsilon\begin{pmatrix} a & b \\ c & d\end{pmatrix} = \begin{cases} \mathrm{e}(-b/24), & c=0,\ d=1 \\ \mathrm{e}(-(a+d)/24c + s(d,c)/2 + 1/8), & c>0 \end{cases} \tag{A.6}$$

where $s(d,c) = \sum_{m=1}^{c-1}(d/c)((md/c))$ and $((x))$ is 0 for $x \in \mathbb{Z}$ and $x - \lfloor x \rfloor - 1/2$ otherwise. We can deduce the values $\epsilon(a,b,c,d)$ for $c < 0$, or for $c = 0$ and $d = -1$, by observing that $\epsilon(-\gamma) = \epsilon(\gamma)\,\mathrm{e}(1/4)$ for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Let T denote the element of $\mathrm{SL}_2(\mathbb{Z})$ such that $\mathrm{tr}(T) = 2$ and $T\tau = \tau + 1$ for $\tau \in \mathbb{H}$. Observe that

$$\epsilon(T^m\gamma) = \epsilon(\gamma T^m) = \mathrm{e}(-m/24)\epsilon(\gamma)$$

for $m \in \mathbb{Z}$.

B Bessel Function

The *Bessel function of the first kind*, denoted $J_\alpha(x)$, may be defined by the series expression

$$J_\alpha(x) = \sum_{m\geq 0}(-1)^m \frac{(x/2)^{2m+\alpha}}{\Gamma(m+1+\alpha)m!}. \tag{B.1}$$

The function $J_\alpha(x)$ is asymptotic to $e^x \cos(x - \alpha\pi/4 - \pi/4)/\sqrt{2\pi x}$ for x large, and in the special case that $\alpha = 1/2$ this fact is borne out by the identity

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin(x). \tag{B.2}$$

The *modified Bessel function of the first kind*, denoted $I_\alpha(x)$, may be defined by the series expression

$$I_\alpha(x) = \sum_{m\geq 0}\frac{(x/2)^{2m+\alpha}}{\Gamma(m+1+\alpha)m!}. \tag{B.3}$$

The function $I_\alpha(x)$ is asymptotic to $e^x/\sqrt{2\pi x}$ for x large, for any α. In the special case that $\alpha = 1/2$ this fact is borne out by the identity

$$I_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sinh(x). \tag{B.4}$$

As x approaches 0 the functions $J_{1/2}(x)$ and $I_{1/2}(x)$ both tend towards $(x/2)^\alpha/\Gamma(1+\alpha)$. Since $\Gamma(3/2)=\sqrt{\pi}/2$ we have

$$J_{1/2}(x)\approx\sqrt{\frac{2x}{\pi}},\ I_{1/2}(x)\approx\sqrt{\frac{2x}{\pi}} \tag{B.5}$$

as $x\to 0$ in the special case that $\alpha=1/2$.

C Lipschitz Summation

The *Lipschitz summation formula* is the identity

$$\frac{(-2\pi i)^s}{\Gamma(s)}\sum_{k=1}^{\infty}(k-\alpha)^{s-1}\,\mathrm{e}((k-\alpha)\tau)=\sum_{\ell\in\mathbb{Z}}\mathrm{e}(\alpha\ell)(\tau+\ell)^{-s}, \tag{C.1}$$

valid for $\Re(s)>1$ and $0\le\alpha<1$, where $\mathrm{e}(x)=e^{2\pi ix}$. A nice proof of this using Poisson summation appears in [46]. Observe that both sides of (C.1) converge absolutely and uniformly in τ on compact subsets of $\mathbb{H}$.

For applications to Rademacher sums of weight less than 1 we require an extension of (C.1) to $s=1$. Absolute convergence on the right hand side breaks down at this point but we can get by with the following useful analogue.

Lemma C.1. *For $0<\alpha<1$ we have*

$$\sum_{k=1}^{\infty}\mathrm{e}((k-\alpha)\tau)=\sum_{-K<\ell<K}\mathrm{e}(\alpha\ell)(-2\pi i)^{-1}(\tau+\ell)^{-1}+E_K(\tau) \tag{C.2}$$

where $E_K(\tau)=\mathcal{O}(1/K^2)$, locally uniformly for $\tau\in\mathbb{H}$.

Proof. Our proof follows that of Lemma 4.1 in [31]. For $\tau\in\mathbb{H}$ define a function $f(z)$ by setting

$$f(z)=\frac{1}{2\pi iz}\frac{\mathrm{e}(\alpha(z-\tau))}{(\mathrm{e}(z-\tau)-1)}. \tag{C.3}$$

Then f has poles at $\tau+n$ for each $n\in\mathbb{Z}$ and also has a pole at $z=0$. The residue at $z=0$ is

$$\frac{1}{2\pi i}\frac{\mathrm{e}(-\alpha\tau)}{\mathrm{e}(-\tau)-1}=\frac{1}{2\pi i}\sum_{k=1}^{\infty}\mathrm{e}((k-\alpha)\tau),$$

which is $(2\pi i)^{-1}$ times the left hand side of (C.2), and the residue at $z=\tau+n$ is $\mathrm{e}(\alpha n)(2\pi i)^{-2}(\tau+n)^{-1}$. Suppose $\tau=\sigma+it$. Let K be a positive integer and let $\mathcal{C}$ be the positively oriented boundary of a rectangle with corners $\sigma\pm(K+1/2)\pm iL$ where $L>\Im(\tau)=t$ and $K+1/2>\Re(\tau)=\sigma$. Then by the residue theorem

$$\int_{\mathcal{C}}f(z)\mathrm{d}z=\sum_{k=1}^{\infty}\mathrm{e}((k-\alpha)\tau)-\sum_{-K<\ell<K}\mathrm{e}(\alpha\ell)(-2\pi i)^{-1}(\tau+\ell)^{-1}.$$

On the other hand, the integrals over the horizontal portions of $\mathcal{C}$ tend to 0 as $L \to \infty$ since $|f(z)|$ decays exponentially as $\Im(z) \to \infty$. The residues are independent of L for L sufficiently large so taking the limit as $L \to \infty$ we obtain

$$\sum_{k=1}^{\infty} \mathrm{e}((k-\alpha)\tau) - \sum_{-K<\ell<K} \mathrm{e}(\alpha\ell)(-2\pi i)^{-1}(\tau+\ell)^{-1} = E_K(\tau)$$

where $E_K(\tau)$ is the limit as $L \to \infty$ of the contributions to $\int_{\mathcal{C}} f(z)\mathrm{d}z$ coming from the vertical portions of $\mathcal{C}$.

$$E_K(\tau) = i\int_{-\infty}^{\infty} (f(\sigma+K+1/2+iy) - f(\sigma-K-1/2+iy))\mathrm{d}y$$

For convenience set $Q = K + 1/2$. After replacing y with $y + t$ we obtain

$$E_K(\tau) = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \left(\frac{\mathrm{e}(\alpha Q)}{Q+iy+\tau} - \frac{\mathrm{e}(-\alpha Q)}{-Q+iy+\tau}\right)\frac{\mathrm{e}(\alpha iy)}{1+\mathrm{e}(iy)}\mathrm{d}y \tag{C.4}$$

since $\mathrm{e}(Q) = \mathrm{e}(-Q) = -1$ for K an integer. Observe that the integrand in (C.4) is bounded by C/Q^2, for some constant C, and decays exponentially as $y \to \pm\infty$ for $0 < \alpha < 1$. Also, the constant C holds locally uniformly in τ. We conclude that $E_K(\tau) = \mathcal{O}(1/K^2)$ locally uniformly in τ, as required. □

For $\alpha = 0$ we have the identity

$$\lim_{K\to\infty} \sum_{\ell=-K}^{K} \frac{1}{\tau+\ell} = \pi\cot\pi\tau = -i\pi\big(1 + 2\sum_{m=1}^{\infty} q^m\big)\ .$$

D Some Details on Convergence

In this appendix we will give the steps leading to the expression (5.5), whose convergence we demonstrated in the main text. Using the explicit definition (4.5), (5.1) and again split the sum over d into a double sum, we obtain

$$\begin{aligned}
&\lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} R^0_\gamma(\tau) \\
&= \mathrm{e}(\tfrac{1}{8}) \sum_{\substack{0<c<K\\ n|c}} c^{-3/2} \sum_{\substack{-K^2<d<K^2\\ (c,d)=1}} \mathrm{e}(-\tfrac{\gamma\infty}{8})\psi(\gamma)(\tau+d/c)^{-1} \\
&= \mathrm{e}(\tfrac{1}{8}) \sum_{\substack{0<c<K\\ n|c}} c^{-3/2} \sum_{\substack{0\le d<c\\ (c,d)=1}} \mathrm{e}(-\tfrac{\gamma\infty}{8})\psi(\gamma) \sum_{\substack{\ell\in\mathbb{Z}\\ |d+\ell c|<K^2}} \frac{\mathrm{e}(\frac{\ell}{8})}{(\tau+d/c+\ell)}\ ,
\end{aligned} \tag{D.1}$$

where we have used $\psi(\gamma T) = \mathrm{e}(\frac{1}{8})\psi(\gamma)$ (see Appendix A) in deriving the last line.

Consider the following two expressions

$$\sum_{\substack{\ell\in\mathbb{Z}\\ |d+\ell c|<K^2}} \mathrm{e}(\tfrac{\ell}{8})(\tau+d/c+\ell)^{-1},\quad \sum_{\substack{\ell\in\mathbb{Z}\\ |\ell|<K^2/c}} \mathrm{e}(\tfrac{\ell}{8})(\tau+d/c+\ell)^{-1}\ , \tag{D.2}$$

we will now argue that we can use either of the two in the limit $K \to \infty$.

The left most of (D.2) may include a term where $d + \ell c = x - K^2$ for at most one $0 \leq x < d$ and this will not appear in the right hand expression, and the right most of (D.2) may include a term where $d + \ell c = x + K^2$ for at most one $0 < x \leq d$ and these will not appear in the left hand expression. We conclude that the difference is $\mathcal{O}(c/K^2)$ for sufficiently large K, uniformly in τ, and since $\sum_{0<c<K} c^{1/2}$ is $\mathcal{O}(K^{3/2})$ we have

$$\sum_{\substack{0<c<K\\ n|c}} c^{1/2} \sum_{\substack{0\leq d<c\\ (c,d)=1}} \mathcal{O}(c/K^2) = \mathcal{O}(1/K^{1/2})$$

which vanishes as $K \to \infty$. Therefore we have

$$\lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} R^0_\gamma(\tau) = \mathrm{e}(\tfrac{1}{8}) \lim_{K\to\infty} \sum_{\substack{0<c<K\\ n|c}} c^{-3/2} \sum_{\substack{0\leq d<c\\ (c,d)=1}} \mathrm{e}(-\tfrac{\gamma\infty}{8})\psi(\gamma) \sum_{\substack{\ell\in\mathbb{Z}\\ |d|<K^2/c}} \frac{\mathrm{e}(\frac{\ell}{8})}{(\tau + d/c + \ell)} \ . \tag{D.3}$$

Upon applying the Lipschitz summation formula in the form of Lemma C.1 we obtain

$$\begin{aligned}\lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)^\times_{<K}} R^0_\gamma(\tau) &= 2\pi\, \mathrm{e}(-\tfrac{1}{8}) \lim_{K\to\infty} \sum_{\substack{0<c<K\\ n|c}} c^{-3/2} \sum_{\substack{0\leq d<c\\ (c,d)=1}} \mathrm{e}(-\tfrac{\gamma\infty}{8})\psi(\gamma)\\ &\times \Big(- E_{\lfloor K^2/c\rfloor}(\tau) + \sum_{k=1}^{\infty} \mathrm{e}((k-\tfrac{1}{8})(\tau + \tfrac{d}{c}))\Big) \ .\end{aligned}$$

For K large the error terms $E_{\lfloor K^2/c\rfloor}(\tau)$ are bounded by $E_K(\tau)$ and according to Lemma C.1 we have $E_K = \mathcal{O}(1/K^2)$. It follows that the term involving $-E_{\lfloor K^2/c\rfloor}(\tau)$ is $\mathcal{O}(1/K^{3/2})$ and thus tends to 0 as $K \to \infty$. Observe that the sum over k in the above formula is absolutely convergent, uniformly in τ on compact subsets of $\mathbb{H}$. Moving the summation over k past the others and using the explicit expressions for $\psi(\gamma)$, we readily obtain (5.5).

E Rademacher Sums at Other Cusps

To identify $R_{\Gamma,\rho}$, we need to study its behaviour at the other cusps of $\Gamma = \Gamma_0(n)$. Focusing on a given cusp representative $\kappa = \sigma\infty$, $\sigma \in \mathrm{SL}_2(\mathbb{Z})$, with width v in $\Gamma_0(n)$, we need to study $R_{\Gamma,\rho}\big|_{\epsilon^{-3},1/2}\sigma$. For this purpose, let us introduce

$$R'_\gamma = \epsilon^{-3}(\gamma)\, e(-\tfrac{\gamma\tau}{8})\, \mathrm{reg}(\gamma,\tau)\, \mathrm{jac}^{1/4}(\gamma,\tau) \tag{E.1}$$

$$S'_\gamma = \epsilon^{3}(\gamma)\, e(\tfrac{\gamma\tau}{8})\, \mathrm{jac}^{3/4}(\gamma,\tau) \tag{E.2}$$

analogous to the definitions in §5, so that

$$R_{\Gamma,\rho} = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \rho(\gamma)\, R'_\gamma \quad , \quad S_{\Gamma,\rho} = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}} \rho^{-1}(\gamma)\, S'_\gamma \ .$$

A straightforward calculation similar to the one in §7 shows

$$\begin{aligned} R'_\gamma\big|_{\epsilon^{-3},1/2}\sigma &= R'_{\gamma\sigma} + cJ_{\sigma^{-1}\infty}S'_{\gamma\sigma} \quad , \quad c = -\frac{\mathrm{e}(\frac{1}{4})}{\sqrt{8}}\\ S'_\gamma\big|_{\epsilon^{3},3/2}\sigma &= S'_{\gamma\sigma} \ . \end{aligned} \tag{E.3}$$

Using $S_{n|1} = \lambda_n \eta^3$ and $S_{n|h} = 0$ for $h > 1$ as shown in §8.1, we get

$$R_{n|1}|_{\epsilon^{-3},1/2}\sigma = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} R'_\gamma + c\lambda_n J_{\sigma^{-1}\infty}\eta^3 \tag{E.4}$$

$$R_{n|h}|_{\epsilon^{-3},1/2}\sigma = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} \rho(\gamma\sigma^{-1})\, R'_\gamma \quad \text{for} \ \ h > 1\ . \tag{E.5}$$

Splitting the sum into a sum over the double coset space $(\Gamma_\infty\backslash\Gamma)_{<K}\sigma/\Gamma^v_\infty$ and applying the Lipschitz summation formula as before, we obtain that the Fourier expansion of the above sum is

$$\lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} \rho(\gamma\sigma^{-1})\, R'_\gamma = \sum_{k=1}^{\infty} c_k\ \mathrm{e}((k-\nu)\tfrac{\tau}{v})\ . \tag{E.6}$$

As before we have $0 < \nu \leq 1$ given by $\nu = \{\{\frac{m}{h} + \frac{v}{8}\}\}$, where $\rho(\sigma T^v \sigma^{-1}) = \mathrm{e}(\frac{m}{h})$ (cf. (7.5)).

In particular, for $h = 1$ we have

$$\lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} R'_\gamma = \sum_{k=1}^{\infty} c_k\ \mathrm{e}(\,(k - \{\{\tfrac{v}{8}\}\}\,)\tfrac{\tau}{v})\ . \tag{E.7}$$

Using Lemma 8.8 and equation (3.6), we can rewrite the function $G_{n|1}$ as

$$G_{n|1} = \lambda_n \tilde{T}_g + \chi(g)\eta^3\big(R_{n|1} + \tfrac{\lambda_n}{12} R_{1|1}\big)\ .$$

From (E.4) we deduce that

$$\big(R_{n|1} + \tfrac{\lambda_n}{12} R_{1|1}\big)\big|_{\epsilon^{-3},1/2}\sigma = \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} R'_\gamma + \tfrac{\lambda_n}{12} R_{1|1}$$

and hence

$$G_{n|1}\big|_{1,2}\,\sigma = \eta^3 \left(-\lambda_n \left(\frac{\chi(g)}{24} H - \frac{\tilde{T}_g|_{1,2}\,\sigma}{\eta^3} \right) + \chi(g) \lim_{K\to\infty} \sum_{\gamma\in(\Gamma_\infty\backslash\Gamma)_{<K}\sigma} R'_\gamma \right)\ .$$

One can check from the explicit expressions for $\tilde{T}_g$ that $\frac{\chi(g)}{24}H - \frac{\tilde{T}_g|_{1,2}\,\sigma}{\eta^3} = \mathcal{O}(1)$, so that the only pole of H_g is at the infinite cusp. Together with the (E.7) we obtain the expansion (8.12).

F Modular Forms and Representations

$1/\eta_g(\tau)$	$1A$	$2A$	$2B$	$3A$	$3B$	$4A$	$4B$	$4C$	$5A$	$6A$	$6B$	$7A$	$7B$	$8A$	$10A$	$11A$	$12A$	$12B$	$14A$	$14B$	$15A$	$15B$	$21A$	$21B$	$23A$	$23B$
q^0	24	8	0	6	0	0	4	0	4	2	0	3	3	2	0	2	0	0	1	1	1	1	0	0	1	1
q^1	324	52	12	27	0	4	16	0	14	7	0	9	9	6	2	5	1	0	3	3	2	2	0	0	2	2
q^2	3200	256	0	104	8	0	48	0	40	16	0	22	22	12	0	10	0	0	4	4	4	4	1	1	3	3
q^3	25650	1122	90	351	0	18	142	6	105	39	0	51	51	28	5	20	3	0	9	9	6	6	0	0	5	5
q^4	176256	4352	0	1080	0	0	368	0	256	80	0	108	108	52	0	36	0	0	12	12	10	10	0	0	7	7
q^5	1073720	15640	520	3107	44	56	928	0	590	175	4	221	221	104	10	65	5	0	23	23	17	17	2	2	11	11
q^6	5930496	52224	0	8424	0	0	2176	0	1296	336	0	432	432	184	0	110	0	0	32	32	24	24	0	0	15	15
q^7	30178575	165087	2535	21762	0	175	4979	27	2740	666	0	819	819	341	20	185	10	0	55	55	37	37	0	0	22	22
q^8	143184000	495872	0	53976	192	0	10864	0	5600	1232	0	1506	1506	580	0	300	0	0	76	76	56	56	3	3	30	30
q^9	639249300	1428612	10908	129141	0	468	23184	0	11130	2289	0	2706	2706	1010	38	481	15	0	122	122	81	81	0	0	42	42

	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	χ_{15}	χ_{16}	χ_{17}	χ_{18}	χ_{19}	χ_{20}	χ_{21}	χ_{22}	χ_{23}	χ_{24}	χ_{25}	χ_{26}
$\mathcal{H}_1$	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\mathcal{H}_2$	3	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\mathcal{H}_3$	6	8	0	0	0	0	3	2	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
$\mathcal{H}_4$	14	20	0	0	0	0	12	6	4	0	0	0	0	1	0	0	3	0	0	0	1	3	0	0	0	0
$\mathcal{H}_5$	27	48	0	0	0	0	33	22	15	0	0	0	0	3	0	0	15	1	0	3	6	16	3	0	0	3
$\mathcal{H}_6$	59	110	0	0	0	0	97	61	51	0	0	0	0	19	0	0	54	10	9	17	34	69	25	6	9	26
$\mathcal{H}_7$	114	249	0	0	6	6	255	174	161	3	3	3	3	70	3	3	190	45	47	88	158	276	147	59	71	194
$\mathcal{H}_8$	235	552	0	0	32	32	687	457	498	40	40	39	39	301	39	39	633	220	269	393	694	1042	758	418	490	1088
$\mathcal{H}_9$	460	1217	1	1	169	169	1783	1235	1504	255	255	296	296	1126	294	294	2152	994	1252	1730	2850	3870	3528	2354	2656	5544
$\mathcal{H}_{10}$	924	2677	24	24	731	731	4754	3294	4575	1425	1425	1675	1675	4329	1699	1699	7207	4391	5592	7131	11460	14340	15393	11758	13026	25565

Table 4: i) The first ten Fourier coefficients of the McKay–Thompson series $Z_g(\tau) = \sum_{k=0}^{\infty} q^{k-1}(\mathrm{Tr}_{\mathcal{H}_k} g) = 1/\eta_g(\tau)$. ii) The decomposition of the first ten M_{24}-modules $\mathcal{H}_k$ into irreducible representations (see the character table E).

$[g]$ ℓ	$1A$	$2A$	$2B$	$3A$	$3B$	$4A$	$4B$	$4C$	$5A$	$6A$	$6B$	$7A$	$7B$	$8A$	$10A$	$11A$	$12A$	$12B$	$14A$	$14B$	$15A$	$15B$	$21A$	$21B$	$23A$	$23B$
0	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
1	90	-6	10	0	6	-6	2	2	0	0	-2	-1	-1	-2	0	2	0	2	1	1	0	0	-1	-1	-2	-2
2	462	14	-18	-6	0	-2	-2	6	2	2	0	0	0	-2	2	0	-2	0	0	0	-1	-1	0	0	2	2
3	1540	-28	20	10	-14	4	-4	-4	0	2	2	0	0	0	0	0	-2	2	0	0	0	0	0	0	-1	-1
4	4554	42	-38	0	12	-6	2	-6	-6	0	4	4	4	-2	2	0	0	0	0	0	0	0	-2	-2	0	0
5	11592	-56	72	-18	0	-8	8	0	2	-2	0	0	0	0	2	-2	-2	0	0	0	2	2	0	0	0	0
6	27830	86	-90	20	-16	6	-2	6	0	-4	0	-2	-2	2	0	0	0	0	2	2	0	0	-2	-2	0	0
7	61686	-138	118	0	30	6	-10	-2	6	0	-2	2	2	-2	-2	-2	0	-2	2	2	0	0	2	2	0	0
8	131100	188	-180	-30	0	-4	4	-12	0	2	0	-3	-3	0	0	2	2	0	-1	-1	0	0	0	0	0	0
9	265650	-238	258	42	-42	-14	10	10	-10	2	6	0	0	-2	-2	0	-2	-2	0	0	2	2	0	0	0	0

χ ℓ	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	χ_7	χ_8	χ_9	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	χ_{15}	χ_{16}	χ_{17}	χ_{18}	χ_{19}	χ_{20}	χ_{21}	χ_{22}	χ_{23}	χ_{24}	χ_{25}	χ_{26}
0	-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	2	2	2	2
8	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	2	0	0	2	2	2	4	2	2	6
9	0	0	0	0	0	0	0	0	2	2	2	0	0	2	2	2	0	2	2	2	4	4	4	8	8	10

Table 5: The first few Fourier coefficients of the terms $q^{-\frac{1}{8}+n}$ in the q-series $H_g(\tau)$ and the corresponding representation K_n (cf. Table E).

classes	$1A$	$2A$	$2B$	$3A$	$3B$	$4A$	$4B$	$4C$	$5A$	$6A$	$6B$	$7A$	$7B$	$8A$	$10A$	$11A$	$12A$	$12B$	$14A$	$14B$	$15A$	$15B$	$21A$	$21B$	$23A$	$23B$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	23	7	-1	5	-1	-1	3	-1	3	1	-1	2	2	1	-1	1	-1	-1	0	0	0	0	-1	-1	0	0
χ_3	45	- 3	5	0	3	- 3	1	1	0	0	- 1	e_7	$\bar{e}_7$	- 1	0	1	0	1	- e_7	- $\bar{e}_7$	0	0	e_7	$\bar{e}_7$	- 1	- 1
χ_4	45	-3	5	0	3	-3	1	1	0	0	-1	$\bar{e}_7$	e_7	-1	0	1	0	1	-$\bar{e}_7$	-e_7	0	0	$\bar{e}_7$	e_7	-1	-1
χ_5	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	e_{15}	$\bar{e}_{15}$	0	0	1	1
χ_6	231	7	-9	-3	0	-1	-1	3	1	1	0	0	0	-1	1	0	-1	0	0	0	$\bar{e}_{15}$	e_{15}	0	0	1	1
χ_7	252	28	12	9	0	4	4	0	2	1	0	0	0	0	2	-1	1	0	0	0	-1	-1	0	0	-1	-1
χ_8	253	13	-11	10	1	-3	1	1	3	-2	1	1	1	-1	-1	0	0	1	-1	-1	0	0	1	1	0	0
χ_9	483	35	3	6	0	3	3	3	-2	2	0	0	0	-1	-2	-1	0	0	0	0	1	1	0	0	0	0
χ_{10}	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	e_{23}	$\bar{e}_{23}$
χ_{11}	770	-14	10	5	-7	2	-2	-2	0	1	1	0	0	0	0	0	-1	1	0	0	0	0	0	0	$\bar{e}_{23}$	e_{23}
χ_{12}	990	-18	-10	0	3	6	2	-2	0	0	-1	e_7	$\bar{e}_7$	0	0	0	0	1	e_7	$\bar{e}_7$	0	0	e_7	$\bar{e}_7$	1	1
χ_{13}	990	-18	-10	0	3	6	2	-2	0	0	-1	$\bar{e}_7$	e_7	0	0	0	0	1	$\bar{e}_7$	e_7	0	0	$\bar{e}_7$	e_7	1	1
χ_{14}	1035	27	35	0	6	3	-1	3	0	0	2	-1	-1	1	0	1	0	0	-1	-1	0	0	-1	-1	0	0
χ_{15}	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2e_7$	$2\bar{e}_7$	-1	0	1	0	-1	0	0	0	0	-e_7	-$\bar{e}_7$	0	0
χ_{16}	1035	-21	-5	0	-3	3	3	-1	0	0	1	$2\bar{e}_7$	$2e_7$	-1	0	1	0	-1	0	0	0	0	-$\bar{e}_7$	-e_7	0	0
χ_{17}	1265	49	-15	5	8	-7	1	-3	0	1	0	-2	-2	1	0	0	-1	0	0	0	0	0	1	1	0	0
χ_{18}	1771	-21	11	16	7	3	-5	-1	1	0	-1	0	0	-1	1	0	0	-1	0	0	1	1	0	0	0	0
χ_{19}	2024	8	24	-1	8	8	0	0	-1	-1	0	1	1	0	-1	0	-1	0	1	1	-1	-1	1	1	0	0
χ_{20}	2277	21	-19	0	6	-3	1	-3	-3	0	2	2	2	-1	1	0	0	0	0	0	0	0	-1	-1	0	0
χ_{21}	3312	48	16	0	-6	0	0	0	-3	0	-2	1	1	0	1	1	0	0	-1	-1	0	0	1	1	0	0
χ_{22}	3520	64	0	10	-8	0	0	0	0	-2	0	-1	-1	0	0	0	0	0	1	1	0	0	-1	-1	1	1
χ_{23}	5313	49	9	-15	0	1	-3	-3	3	1	0	0	0	-1	-1	0	1	0	0	0	0	0	0	0	0	0
χ_{24}	5544	-56	24	9	0	-8	0	0	-1	1	0	0	0	0	-1	0	1	0	0	0	-1	-1	0	0	1	1
χ_{25}	5796	-28	36	-9	0	-4	4	0	1	-1	0	0	0	0	1	-1	-1	0	0	0	1	1	0	0	0	0
χ_{26}	10395	-21	-45	0	0	3	-1	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	-1	-1

Table 6: Character table of M_{24}. See [23]. We adopt the naming system of [23] and use the notation $e_n = \frac{1}{2}(-1 + \sqrt{-n})$.

References

[1] T. Eguchi, H. Ooguri, and Y. Tachikawa, "Notes on the K3 Surface and the Mathieu group M_{24}," `1004.0956`.

[2] M. C. N. Cheng, "K3 Surfaces, N=4 Dyons, and the Mathieu Group M_{24}," `1005.5415`.

[3] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu twining characters for K3," `1006.0221`.

[4] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Mathieu Moonshine in the elliptic genus of K3," `1008.3778`.

[5] T. Eguchi and K. Hikami, "Note on Twisted Elliptic Genus of K3 Surface," `1008.4924`.

[6] J. H. Conway and S. P. Norton, "Monstrous Moonshine," *Bull. London Math. Soc.* **11** (1979) 308 339.

[7] R. E. Borcherds, "Monstrous moonshine and monstrous Lie superalgebras," *Invent. Math.* **109, No.2** (1992) 405–444.

[8] S. Norton, "Generalized Moonshine," *Proc. Symp. Pure Math* **47** (1987) 208–209.

[9] A. P. Ogg, "Automorphismes des Courbes Modulaires," *Seminaire Delange-Pisot* (1974) .

[10] J. F. R. Duncan and I. B. Frenkel, "Rademacher sums, moonshine and gravity," `0907.4529`.

[11] H. Rademacher, "The Fourier Series and the Functional Equation of the Absolute Modular Invariant J(τ)," *Amer. J. Math.* **61** (1939) no. 1, 237–248.

[12] J. M. Maldacena, "The Large N Limit of Superconformal Field Theories and Supergravity," *Adv.Theor.Math.Phys.* **2:231-252,1998** (Adv.Theor.Math.Phys.2:231-252,1998) , `hep-th/9711200`.

[13] J. Maldacena and A. Strominger, "AdS3 Black Holes and a Stringy Exclusion Principle," *JHEP* **9812:005,1998** (JHEP 9812:005,1998) , `hep-th/9804085`.

[14] R. Dijkgraaf, J. Maldacena, G. Moore, and E. Verlinde, "A Black Hole Farey Tail," `hep-th/0005003`.

[15] G. W. Moore, "Les Houches Lectures on Strings and Arithmetic," `hep-th/0401049`.

[16] J. de Boer, M. C. N. Cheng, R. Dijkgraaf, J. Manschot, and E. Verlinde, "A Farey Tail for Attractor Black Holes," *JHEP* **0611:024,2006** (JHEP0611:024,2006) , `hep-th/0608059`.

[17] P. Kraus and F. Larsen, "Partition functions and elliptic genera from supergravity," *JHEP* **0701:002,2007** (JHEP 0701:002,2007) , `hep-th/0607138`.

[18] F. Denef and G. W. Moore, "Split States, Entropy Enigmas, Holes and Halos," `hep-th/0702146`.

[19] J. Manschot and G. W. Moore, "A Modern Fareytail,"*Commun.Num.Theor.Phys.* **4,** (Dec., 2007) 103–159,2010, `0712.0573`.

[20] T. Eguchi and K. Hikami, "Superconformal Algebras and Mock Theta Functions 2. Rademacher Expansion for K3 Surface," *Communications in Number Theory and Physics* **3,** (Apr., 2009) 531–554, 0904.0911.

[21] W. d. A. Pribitkin, "The Fourier coefficients of modular forms and Niebur modular integrals having small positive weight. II," *Acta Arith.* **93** (2000) no. 4, 343–358.

[22] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups.* Springer-Verlag, 1999.

[23] J. Conway, R. Curtis, S. Norton, R. Parker, and R. Wilson, *Atlas of finite groups. Maximal subgroups and ordinary characters for simple groups. With comput. assist. from J. G. Thackray.* Oxford: Clarendon Press, 1985.

[24] G. Mason, "M_{24} and Certain Automorphic Forms," *Contemporary Mathematics* **45** (1985) 223–244.

[25] D. Dummit, H. Kisilevsky, and J. McKay, "Multiplicative products of eta-functions.," *Contemp. Math. (Finite groups - coming of age, Proc. CMS Conf., Montreal/Que. 1982)* **45** (1985) .

[26] S. Govindarajan and K. G. Krishna, "BKM Lie superalgebras from dyon spectra in Z_N-CHL orbifolds for composite N," 0907.1410.

[27] M. C. N. Cheng and J. F. R. Duncan, "Automorphic forms and the largest Mathieu group," *To appear* .

[28] T. Eguchi, H. Ooguri, A. Taormina, and S.-K. Yang, "Superconformal Algebras and String Compactification on Manifolds with SU(N) Holonomy," *Nucl. Phys.* **B315** (1989) 193.

[29] T. Eguchi and K. Hikami, "Superconformal algebras and mock theta functions," *J.Phys.A* **42:304010,2009** (Dec., 2008) , 0812.1151.

[30] A. Dabholkar, S. Murthy, and D. Zagier, "Quantum Black Hole, Wall-Crossing, and Mock Modular Forms," *to appear* .

[31] D. Niebur, "Construction of automorphic forms and integrals," *Trans. Amer. Math. Soc.* **191** (1974) 373–385.

[32] D. Goldfeld and P. Sarnak, "Sums of Kloosterman sums," *Invent. Math.* **71** (1983) no. 2, 243–250. `http://dx.doi.org/10.1007/BF01389098`.

[33] A. Selberg, "On the estimation of Fourier coefficients of modular forms," in *Proc. Sympos. Pure Math., Vol. VIII*, pp. 1–15. Amer. Math. Soc., Providence, R.I., 1965.

[34] W. Roelcke, "Das Eigenwertproblem der automorphen Formen in der hyperbolischen Ebene. I, II," *Math. Ann. 167 (1966), 292–337; ibid.* **168** (1966) 261–324.

[35] P. Sarnak, "Additive number theory and Maass forms," in *Number theory (New York, 1982)*, vol. 1052 of *Lecture Notes in Math.*, pp. 286–309. Springer, Berlin, 1984. `http://dx.doi.org/10.1007/BFb0071548`.

[36] M. I. Knopp, "On the Fourier coefficients of small positive powers of $\theta(\tau)$," *Invent. Math.* **85** (1986) no. 1, 165–183. `http://dx.doi.org/10.1007/BF01388796`.

[37] M. I. Knopp, "On the Fourier coefficients of cusp forms having small positive weight," in *Theta functions—Bowdoin 1987, Part 2 (Brunswick, ME, 1987)*, vol. 49 of *Proc. Sympos. Pure Math.*, pp. 111–127. Amer. Math. Soc., Providence, RI, 1989.

[38] H. Davenport, *Multiplicative number theory*, vol. 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third ed., 2000. Revised and with a preface by Hugh L. Montgomery.

[39] W. d. A. Pribitkin, "A generalization of the Goldfeld-Sarnak estimate on Selberg's Kloosterman zeta-function," *Forum Math.* **12** (2000) no. 4, 449–459. `http://dx.doi.org/10.1515/form.2000.014`.

[40] G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 18. Stechert-Hafner, Inc., New York, 1964.

[41] C. Vafa, "Gas of D-Branes and Hagedorn Density of BPS States," *Nucl.Phys. B* **463** (1996) 415–419.

[42] A. Dabholkar, J. Gomes, and S. Murthy, "Quantum black holes, localization and the topological string," *JHEP* **06** (2011) 019, `arXiv:1012.0265 [hep-th]`.

[43] A. Dabholkar, J. Gomes, and S. Murthy, "Localization and Exact Holography," *To appear* .

[44] E. Witten, "Three-Dimensional Gravity Revisited," `0706.3359`.

[45] M. R. Gaberdiel, S. Hohenegger, and R. Volpato, "Symmetries of K3 sigma models," `1106.4315`.

[46] M. Knopp and S. Robins, "Easy proofs of Riemann's functional equation for $\zeta(s)$ and of Lipschitz summation," *Proc. Amer. Math. Soc.* **129** (2001) no. 7, 1915–1922 (electronic).

AEI-2013-269

arXiv:1312.0622v3 [hep-th] 11 Nov 2014

SECOND QUANTIZED MATHIEU MOONSHINE

Daniel Persson
Fundamental Physics, Chalmers University of Technology,
412 96, Gothenburg, Sweden
`daniel.persson@chalmers.se`

Roberto Volpato
Max-Planck-Institut für Gravitationsphysik,
Am Mühlenberg 1, 14476 Golm, Germany
`roberto.volpato@aei.mpg.de`

Abstract

We study the second quantized version of the twisted twining genera of generalized Mathieu moonshine, and prove that they give rise to Siegel modular forms with infinite product representations. Most of these forms are expected to have an interpretation as twisted partition functions counting 1/4 BPS dyons in type II superstring theory on K3$\times T^2$ or in heterotic CHL-models. We show that all these Siegel modular forms, independently of their possible physical interpretation, satisfy an "S-duality" transformation and a "wall-crossing formula". The latter reproduces all the eta-products of an older version of generalized Mathieu moonshine proposed by Mason in the '90s. Surprisingly, some of the Siegel modular forms we find coincide with the multiplicative (Borcherds) lifts of Jacobi forms in umbral moonshine.

Contents

1 Introduction and Summary **3**

1.1 Background 4

1.2 Summary of results 6

1.3 Outline 9

2 Twisted twining partition functions **10**

2.1 Definition and basic properties 10

2.2 Modular properties and the α-twist 12

2.3 (Generalized) Mathieu moonshine 14

2.4 Geometric perspective 16

3 Second quantization of Mathieu moonshine **17**

3.1 Symmetric Orbifolds and Hecke Operators 17

3.2 Twisted equivariant Hecke operators 19

3.2.1 Generalities on Hecke operators 19

3.2.2 Twisted equivariant Hecke operators: main definition 20

3.2.3 Central extensions 22

3.2.4 Twisted equivariant Hecke operators: alternative definition 23

3.3 Second-Quantized Twisted Twining Genera 23

3.3.1 Definition 24

3.3.2 Infinite product representation 25

3.3.3 Multiplicative versus additive lift 28

4 Wall-crossing and Mason's generalized moonshine for M_{24} **29**

4.1 Mason's generalized M_{24}-moonshine 30

4.2 Connecting the two moonshines via wall-crossing 31

5 Automorphic Properties **34**

5.1 Relabeling and S-duality 34

5.1.1 Orbifolds and relabeling 34

5.1.2 Relabeling for Mathieu Moonshine 37

5.1.3 S-duality invariance 40

5.2 Automorphic properties 42

5.3 Examples 46

6 Conclusions **51**

A Some group cohomology 53

B Modular properties of twisted twining partition functions 56
B.1 $SL(2,\mathbb{Z}) \times G$ - action on $\mathcal{P}_G$ 56
B.2 Twisted action on $\mathcal{P}_G$: generalized permutation 56
B.3 Modular properties of twisted twining partition functions 57
B.4 Cohomological obstructions 58

C Definition of twisted equivariant Hecke operators 59
C.1 Twisted equivariant Hecke operator 59
C.2 Definition of slash operator for $\mathrm{Mat}(\mathbb{Z}) \times G$ 61
C.3 Proof of consistency 63

D Some technical proofs 65
D.1 Proof of (3.47). 65
D.2 Central extensions and special choices of the cocycle 67
D.2.1 Case $\lambda = 1$ (trivial multiplier) 68
D.2.2 Case $\lambda \neq 1$ (non-trivial multiplier) 69
D.3 S-duality for g in classes $2B, 3B, 4C, 6B, 12B$ 70

E Modular forms and Jacobi forms 75

F Tables 76

1 Introduction and Summary

Mathieu moonshine [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] and its generalized version [17, 18] pertain to the association of a class of weak Jacobi forms $\phi_{g,h}(\tau, z)$, called twisted twining genera, to each commuting pair of elements (g, h) in M_{24}. It has been verified that these functions satisfy all the requirements of Norton's generalized moonshine conjectures [19]; in particular they decompose into (projective) graded characters of the centralizer of g in M_{24}. Many of these Jacobi forms arise as supersymmetric indices in certain non-linear sigma models with target space K3; in particular, $\phi_{e,e}$ is the K3 elliptic genus. The obvious idea that M_{24} might be the symmetry group of some of these non-linear sigma models has been ruled out quite quickly [8, 10, 11] and the most recent works on the subject [20, 21, 22] seek for an explanation of Mathieu moonshine in the framework of K3 compactifications of full-fledged superstring theory rather than within conformal field theory.

In this paper we adopt a similar viewpoint, and construct the second-quantized[*] version $\Psi_{g,h}$ of the twisted twining genera. From a mathematical perspective, the functions $\Psi_{g,h}$ are obtained from the Jacobi forms $\phi_{g,h}$ via a certain twisted equivariant version of Borcherds multiplicative lift. Physically, many of these functions can be interpreted as twisted supersymmetric indices counting 1/4 BPS states in type II superstring theory compactified on K3$\times T^2$ and in CHL models [24, 25, 26, 27]. A similar construction for the twining genera of the ordinary Mathieu moonshine was first proposed by Cheng [2].

We show that these multiplicative lifts are Siegel modular forms, generalizing results for the case $(g, h) = (e, h)$ [2, 28, 29]. In some cases we can identify $\Psi_{g,h}$ with known Siegel modular forms which have previously appeared in the context of umbral moonshine [30]. Furthermore, they satisfy a "wall-crossing formula", which reproduces all of Mason's generalized eta-products $\eta_{g,h}$ for M_{24} [31, 32, 33, 34]. This therefore establishes the link between the two existing versions of generalized moonshine for M_{24}.

1.1 Background

The first hint of Mathieu moonshine was found by Eguchi, Ooguri and Tachikawa [1], who noticed a connection between the the elliptic genus of K3 and the largest Mathieu group M_{24}. The elliptic genus of K3 $\phi_{K3}(\tau, z)$ is the unique (up to normalization) weak Jacobi form of weight 0 and index 1 and it can be defined as a supersymmetric index of non-linear sigma models with target space K3. It therefore is natural to consider a decomposition of $\phi_{K3}(\tau, z)$ into characters of the $\mathcal{N} = 4$ superconformal algebra. The authors of [1] noticed that the coefficients in this decomposition are sums of dimensions of irreducible representations of M_{24}. This observation led to the Mathieu moonshine conjecture: for each conjugacy class $[g]$ of M_{24} there should exist a weak Jacobi form $\phi_g : \mathbb{H}\times\mathbb{C} \to \mathbb{C}$ such that the Fourier coefficients of ϕ_g are the characters of these M_{24}-representations evaluated at g. In particular, for the identity element e of M_{24}, $\phi_e(\tau, z)$ coincides with the elliptic genus of K3. In subsequent work [2, 3, 4, 5], all the functions ϕ_g, dubbed "twining elliptic genera", were found and substantial evidence was given for the validity of Mathieu moonshine, namely the existence of a graded M_{24}-module such that the ϕ_g's are its graded characters. This conjecture has now been proven rigorously by Gannon [12].

The story outlined above is of course in close analogy with Monstrous moonshine [35, 36, 37] which assigns modular functions $T_g : \mathbb{H} \to \mathbb{C}$ (McKay-Thompson series) with each conjugacy class $[g]$ of the Fischer-Griess Monster group $\mathbb{M}$, the largest of the finite sporadic simple groups. After the initial conjecture [35], Norton proposed his generalized Monstrous moonshine [19], which assigns modular functions $f(g, h; \tau)$ to each commuting pair $g, h \in \mathbb{M}$. For fixed $g \in \mathbb{M}$, these generalized moonshine functions should then have (possibly rational) Fourier coefficients that correspond to projective characters of the centralizer $C_{\mathbb{M}}(g) = \{k \in \mathbb{M} \,|\, gk = kg\}$. Dixon-Ginsparg-Harvey subsequently suggested that Norton's functions $f(g, h; \tau)$ naturally arise in string theory as the path integral on a torus $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ with boundary conditions twisted by (g, h) along the (a, b)-cycles. Although the full generalized Monstrous moonshine conjecture is still open, considerable progress has been made toward proving it [38, 39, 40, 41, 42, 43, 44].

---

[*] The terminology "second-quantized" originates in the work of Dijkgraaf, Moore, Verlinde, Verlinde [23].

In earlier work [17, 18], we gave substantial evidence that Norton's generalization also holds, with small modifications, for Mathieu moonshine. We found that for each commuting pair $g, h \in M_{24}$ there exists a weak Jacobi form $\phi_{g,h}(\tau, z)$, dubbed *twisted twining genus*, whose Fourier coefficients are characters of a projective representation of $C_{M_{24}}(g)$. Inspired by orbifolds of holomorphic CFT's [45, 46, 47, 48], we further showed that the modular properties of the functions $\phi_{g,h}$ are controlled by a cohomology class $[\alpha]$ in the third cohomology group $H^3(M_{24}, U(1))$, as was anticipated in [49] (see also [12]).

A different kind of generalized moonshine for M_{24} had in fact already been established in old work by Mason [31, 33, 34]. Mason associated to each commuting pair g, h in M_{24} a so called multiplicative eta product $\eta_{g,h}(\tau)$, based on the action of M_{24} on 24 chiral free bosons. This leads to the natural question: *Is there a relation between the recently discovered Mathieu moonshine, pertaining to weak Jacobi forms, and Mason's M_{24}-moonshine involving eta-products?*

For the special case of commuting pairs $(g, h) = (e, h)$, where e is the identity element of M_{24}, this was given an affirmative answer by Cheng in [2]. Cheng's idea was to generalize the known fact that the elliptic genus of K3, $\phi_e(\tau, z)$, exhibits an exponential Borcherds lift to the unique weight 10 Siegel modular form for $Sp(4; \mathbb{Z})$, referred to as the "Igusa cusp form" and commonly denoted by Φ_{10}. She proposed that to all twining genera $\phi_h(\tau, z)$ one should have a similar lift of the form

$$\Phi_h(\sigma, \tau, z) = pqy \prod_{(n,m,\ell)>0} \exp\left[- \sum_{k=1}^{\infty} \frac{c_{h^k}(4mn - \ell^2)}{k} (p^m q^n y^\ell)^k \right], \tag{1.1}$$

where $q = e^{2\pi i\tau}, y = e^{2\pi i z}, p = e^{2\pi i\sigma}$, and c_g denotes the Fourier coefficients of $\phi_g(\tau, z)$:

$$\phi_g(\tau, z) = \sum_{m\geq 0, \ell\in\mathbb{Z}} c_g(m, \ell) q^m y^\ell. \tag{1.2}$$

It was conjectured in [28] that Φ_h is automorphic with respect to a subgroup $\Gamma_h^{(2)}$ of $Sp(4; \mathbb{Z})$. This was then proven by Raum [29] for most of the conjugacy classes $[h] \subset M_{24}$.

A central point for us is that $\Phi_h(\sigma,\tau,z)$ has a double pole at $z=0$ and in the limit one finds

$$\lim_{z\to 0}\frac{\Phi_h(\sigma,\tau,z)}{(2\pi i z)^2}=\eta_h(\tau)\,\eta_h(\sigma), \tag{1.3}$$

where $\eta_h=\eta_{e,h}$ are the eta-products of Mason's M_{24}-moonshine [31]. Hence, the process of taking the multiplicative lift of ϕ_h followed by studying the limiting behavior as $z\to 0$ provides a link between the two moonshines.

1.2 Summary of results

In this paper we answer the question above in the general case of commuting pairs (g,h) in M_{24}. In other words, we establish a link between the generalized Mathieu moonshine proposed in [17, 18] and the generalized eta-products of Mason. To describe our results, we recall the notion of "second quantized elliptic genus" as defined in [23]. Suppose $\phi_X(\tau,z)$ is the elliptic genus of some Calabi-Yau manifold X. Then one defines the second quantized genus Ψ_X as the generating function of the elliptic genus ϕ_{S^nX} for the n:th symmetric product S^nX: $\Psi_X=\sum_{n\geq 0}p^n\phi_{S^nX}$. In [23], the following remarkable formula is proved

$$\Psi_X(\sigma,\tau,z)=\exp\Big[\sum_{n=1}^{\infty}p^nT_n\phi_X(\tau,z)\Big], \tag{1.4}$$

where $p=e^{2\pi i\sigma}$, and T_n is the standard Hecke operator acting on Jacobi forms. Subsequently, Gritsenko showed [50] that by multiplying the inverse Ψ_X^{-1} by a certain factor $A_X(\sigma,\tau,z)$ (the "Hodge anomaly") one obtains a Siegel modular form

$$\Phi_X(\sigma,\tau,z)=\frac{A_X(\sigma,\tau,z)}{\Psi_X(\sigma,\tau,z)}, \tag{1.5}$$

of weight $c_X(0,0)/2$ for a subgroup $\Gamma_X^{(2)}\subset Sp(4;\mathbb{Z})$, where $c_X(m,n)$ are the Fourier coefficients of ϕ_X.

Inspired by these results, we define for each commuting pair (g,h) in M_{24} the associated *second quantized twisted twining genus* $\Psi_{g,h}$ via a generalization of the formula (1.4):

$$\Psi_{g,h}(\sigma,\tau,z)=\exp\Big[\sum_{n=1}^{\infty}p^n\mathcal{T}_n^{\alpha}\phi_{g,h}(\tau,z)\Big], \tag{1.6}$$

where $\mathcal{T}_n^{\alpha}$ is now a certain twisted equivariant Hecke operator (for the precise definition see section (3.2)) which reduces to the T_n in (1.4) in the special case $(g,h)=(e,e)$. Notice that, when $g=e$, the presence of a non trivial 3-cocycle α governing the modular properties of the twisted twining genera can be safely ignored (see [17] and section 2.2). This leads to the simpler definition of the multiplicative lift adopted in [28]. However, for the general case considered in this paper, the technical subtleties associated with non-trivial 3-cocycles cannot be avoided.

Multiplying the second quantized twisted twining genera by a correction $A_{g,h}(\sigma,\tau,z)$ (see eq. (3.53) for the explicit form of $A_{g,h}$) we obtain a new class of functions

$$\Phi_{g,h}(\sigma,\tau,z)=\frac{A_{g,h}(\sigma,\tau,z)}{\Psi_{g,h}(\sigma,\tau,z)}, \tag{1.7}$$

and the main result of this paper is:

Theorem. *The functions $\Phi_{g,h}(\sigma,\tau,z)$ defined by (1.7) are meromorphic Siegel modular forms for certain discrete subgroups $\Gamma^{(2)}_{g,h}\subset Sp(4;\mathbb{R})$, such that $\Gamma^{(2)}_{g,h}\cap Sp(4,\mathbb{Z})$ is a congruence subgroup of $Sp(4,\mathbb{Z})$.*

For a more precise formulation, see Theorem 5.3 in section 5.2.

The most noticeable new automorphic property is the 'S-duality' transformation

$$\Phi_{g,h}(\sigma,\tau,z)=\Phi_{g,h'}(\frac{\tau}{N\lambda},N\lambda\sigma,z)\ , \tag{1.8}$$

where N is the order of g and λ is the length of the shortest cycle of g in the 24-dimensional permutation representation. Notice that $h,h'\in C_{M_{24}}(g)$ are not necessarily in the same conjugacy class. This is related with the 'relabeling' phenomenon described in [17]. In particular, as described in section 5.1.2, when g is an element of an M_{23} subgroup, then h' is the symmetry induced by h in the g-orbifold theory. We proved (1.8) for at least one commuting pair of generators (g,h) in each group $\langle g,h\rangle$, but we conjecture that similar relations hold for all commuting pairs.

In section 4 we prove the following "wall-crossing formula"

$$\lim_{z\to 0}\frac{\Phi_{g,h}(\sigma,\tau,z)}{(2\pi i z)^2}=\eta_{g,h}(\tau)\eta_{g,h'}(N\lambda\sigma), \tag{1.9}$$

which generalizes (1.3). This shows that the limit $z\to 0$ of $\Phi_{g,h}(\sigma,\tau,z)$ reproduces all of Mason's generalized eta-products $\eta_{g,h}$ [33, 32], thus providing the desired link between the two moonshines. See figure 1.2 for a pictorial overview of the relation between the various modular objects.

When (g,h) lies in the conjugacy classes $(2A,2A),(3A,3A)$ and $(4B,4B)$ we can identify $\Phi_{g,h}$ with known Siegel modular forms:

$$\begin{aligned}(2A,2A)\ &:\ \Phi_{g,h}=(\Delta_2)^2 &&=\Phi^{(3)},\\ (3A,3A)\ &:\ \Phi_{g,h}=(\Delta_1)^2 &&=\Phi^{(4)},\\ (4B,4B)\ &:\ \Phi_{g,h}=(\Delta_{1/2})^2 &&=\Phi^{(5)},\end{aligned} \tag{1.10}$$

where the first equality identifies them with the squares of the weight k Siegel modular forms Δ_k obtained by Gritsenko and Nikulin [51], and the second equality with the umbral Siegel modular forms $\Phi^{(\ell)}$ of lambency $\ell = 3, 4, 5$ [30]. Thus we obtain a surprising connection between generalized Mathieu moonshine and umbral moonshine, which clearly deserves further investigation.[†]

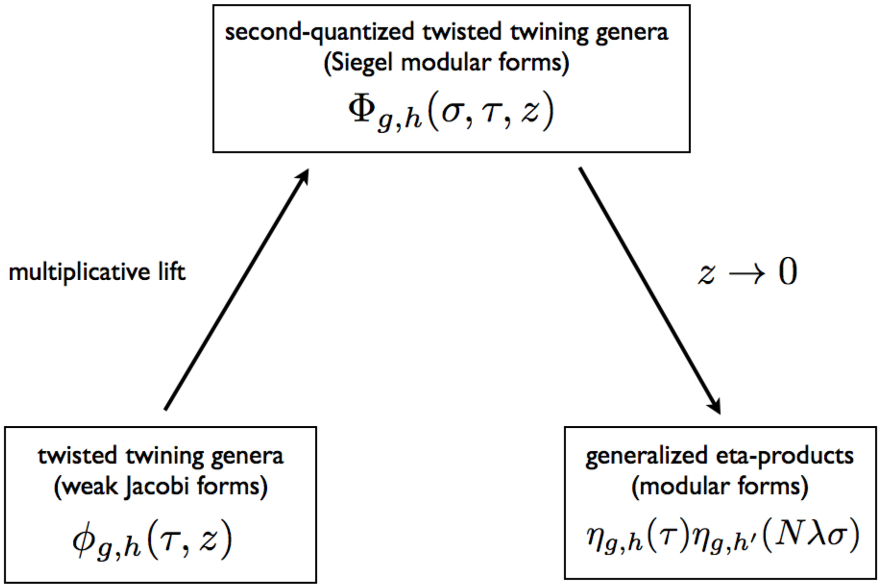

second-quantized twisted twining genera
(Siegel modular forms)
$\Phi_{g,h}(\sigma,\tau,z)$
multiplicative lift
$z \to 0$
twisted twining genera
(weak Jacobi forms)
$\phi_{g,h}(\tau,z)$
generalized eta-products
(modular forms)
$\eta_{g,h}(\tau)\eta_{g,h'}(N\lambda\sigma)$

Figure 1: Pictorial description of the relation between generalized Mathieu moonshine, involving the twisted twining genera $\phi_{g,h}$, and Mason's generalized eta-products $\eta_{g,h}$. Starting from $\phi_{g,h}$ one constructs a second-quantized twisted twining genus $\Psi_{g,h}$, whose reciprocal becomes a Siegel modular form $\Phi_{g,h}$ after multiplying by a factor $A_{g,h}$. This is the content of the multiplicative (automorphic) lift indicated in the leftmost arrow. At the level of the Siegel modular form $\Phi_{g,h}$ one then takes the limit $z \to 0$ which produces the generalized eta-products $\eta_{g,h}(\tau)\eta_{g,h'}(N\lambda\sigma)$. Here, h' is possibly in a different class from h, as a consequence of the 'relabeling phenomenon' described in [17]. See section 4.2 for more details.

Finally, we would like to comment further on the possible physical interpretation of the functions $\Phi_{g,h}$. For $\Phi_{e,e} = \Phi_{10}$, it is of course well-known that the reciprocal $1/\Phi_{10}$ is the generating function of 1/4-BPS dyons in heterotic string theory on T^6 [52], and in this case the decomposition (1.3) becomes $1/\Phi_{10} \sim \eta(\tau)^{-24}\eta(\rho)^{-24}$, where each factor on the right hand side is identified with the partition function of 1/2 BPS-states. More generally, $1/\Phi_{e,h}$ should correspond to the partition function of "twisted dyons", and the limit (1.3) reflects the wall-crossing phenomenon when a 1/4 BPS-state splits into two 1/2 BPS-states. Similarly we expect that in general $1/\Phi_{g,h}$ should be the generating function of certain "twisted dyons" in CHL orbifolds $T^6/\mathbb{Z}_N$ of heterotic string theory [24, 25, 26, 27]. In particular, this should be true whenever g, h can be interpreted as a pair of commuting symmetries of some K3 surface and the orbifold by g is consistent. Given the results of [8], the first condition amounts to the group $\langle g, h\rangle$ having at least four orbits in the standard 24-dimensional permutation representation of M_{24}, while the second condition simply requires g to belong to some M_{23} subgroup of M_{24}. From the tables in appendix F, it is easy to check that the first condition is satisfied by all 34 groups except the groups numbered 22, 30, 31, 32. As for the remaining groups, the second condition is satisfied by at least one generator g, except for groups 4, 5, 6, 13, 14.

[†] We thank Miranda Cheng for a discussion that pointed us in this direction.

1.3 Outline

Our paper is organized as follows. In section 2 we recall some relevant facts about holomorphic orbifolds, focussing on properties of the associated twisted twinining partition functions. We explain the important role played by the third cohomology group $H^3(G, U(1))$, where G is a finite automorphism group of the holomorphic CFT. In particular, this cohomology group controls the modular properties of the twisted twining partition functions. After discussing these properties in a general context we restrict to the relevant case of $G = M_{24}$ for which we recall the main results of [17, 18]. Section 3 introduces some aspects of symmetric products of holomorphic CFT's and the connection with Hecke operators. In section 3.2 we then introduce the key notion of twisted equivariant Hecke operators that will play a crucial role in what follows. In section 3.3 we use the twisted equivariant Hecke operators to define second-quantized twisted twining genera $\Psi_{g,h}$ and we compute their infinite product expansions. The connection with Mason's eta-products is then made in section 4 by evaluating $\Psi_{g,h}$ in the limit $z \to 0$. The modular properties of the second-quantized twisted twining genera are analyzed in section 5.2. Finally, in the concluding section 6 we summarize our results and give some ideas for future research. Various background material, technical calculations and tables have been relegated to the appendices.

2 Twisted twining partition functions

In this section we review the main properties of the twisted twining partition functions associated with the orbifold of a conformal field theory (CFT) $\mathcal{C}$ by a finite group G of symmetries [53, 54, 39]. When $\mathcal{C}$ is a bosonic CFT, these are modular functions $Z_{g,h}(\tau)$ on the upper-half plane $\mathbb{H}$, associated with each commuting pair $g, h \in G$. For superconformal field theories the analogous functions include the twisted twining genera $\phi_{g,h}(\tau, z)$, which are Jacobi forms on $\mathbb{H} \times \mathbb{C}$. Our main interest is the case when $\mathcal{C}$ is a putative $\mathcal{N} = 4$ super-vertex algebra at central charge $c = 6$ with automorphism group $G = M_{24}$, but we shall keep the discussion general whenever possible. We emphasize that the existence of such a super-vertex algebra is not established and its definition seems to be problematic. Despite these issues, the properties satisfied by the twisted-twining genera are analogous to the ones expected for a non-linear sigma model with target space K3, although no such model has a group of symmetries containing M_{24}. Finally, we will provide an abstract formulation of Mathieu moonshine [1, 2, 3, 4, 5, 12] and of its generalized version [17, 18].

2.1 Definition and basic properties

Consider a two-dimensional conformal field theory $\mathcal{C}$ and let G be its symmetry group, i.e. the group of linear automorphisms of its space of states $\mathcal{H}$ that preserves the OPE of the corresponding fields and fixes (at least) the left- and right-moving Virasoro algebra. Given such data, one can construct a new conformal field theory by considering the orbifold of $\mathcal{C}$ by (a subgroup of) G. The first step in the orbifold construction consists in introducing, for each $g \in G$, the g-twisted fields Ξ_g, generating a g-twisted space of states $\mathcal{H}_g$. The defining property of the g-twisted fields is that every field of $\mathcal{C}$ has a non-trivial monodromy g when it is moved along a (sufficiently small) closed path encircling the twisted field Ξ_g. The orbifold theory is defined by including all twisted sectors in the spectrum and then restricting to the G-invariant part. The orbifold theory is a consistent CFT provided certain conditions (in particular, the level-matching condition), assuring the locality of the OPE of twisted fields, are satisfied.

Even if the consistency conditions for the orbifold theory are not satisfied, it makes sense to consider the g-twisted sector $\mathcal{H}_g$ as a module over the Virasoro algebra (or more generally, over the G-invariant part of the chiral algebra of $\mathcal{C}$). Let us consider the case where $\mathcal{C}$ is a holomorphic CFT, so that it coincides with its chiral algebra, which is a self-dual vertex operator algebra. Then, under suitable assumptions, there exists a unique irreducible g-twisted sector $\mathcal{H}_g$ for each $g \in G$ [39]. The action of a generic element $h \in G$ on $\mathcal{C}$ induces a linear map

$$\rho_g(h) : \mathcal{H}_g \to \mathcal{H}_{h^{-1}gh} \ , \tag{2.1}$$

from the g-twisted to the $h^{-1}gh$-twisted sector. In particular, when h commutes with g, it defines an endomorphism of $\mathcal{H}_g$, so that $\mathcal{H}_g$ carries a representation ρ_g of the centralizer

$$C_G(g) := \{h \in G : gh = hg\} \ , \tag{2.2}$$

of g in G. It should be noticed that the representation ρ_g of $C_G(g)$ is in general only projective. This fact will play a crucial role in what follows.

Let us define a g-twisted h-twining partition function on the upper half-plane $\mathbb{H}_+$ by

$$Z_{g,h}(\tau) = \mathrm{Tr}_{\mathcal{H}_g}(\rho_g(h) q^{L_0 - \frac{c}{24}}) \ , \qquad q := e^{2\pi i \tau}, \quad g, h \in G, \ gh = hg \ . \tag{2.3}$$

When g and h are the identity, this reduces to the usual torus partition function for $\mathcal{C}$. On general physical grounds, one expects the twisted twining partition function $Z_{g,h}$ to be given by a path-integral on an elliptic curve $E_\tau \cong \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ with modular parameter τ, where the fields have monodromies g and h around the generators $-1, \tau$ of the first homology group $H_1(E_\tau, \mathbb{Z}) \cong \mathbb{Z} + \tau\mathbb{Z}$. Furthermore, each twisted twining partition function $Z_{g,h}$ is a holomorphic function of the modular parameter τ and is expected to be a modular function under the subgroup $\Gamma_{g,h} \subset SL(2, \mathbb{Z})$, that preserves the monodromies of the fields around the non-trivial cycles of the torus.

Analogous properties are expected for the twisted twining genera in superconformal field theories with an extended (at least $\mathcal{N} = (2, 2)$) superconformal algebra and a group G of symmetries preserving such an algebra. The twisted twining genera are defined as traces

$$\phi_{g,h}(\tau, z) = \mathrm{Tr}_{\mathcal{H}_g^{RR}}(h q^{L_0 - \frac{c}{24}} \bar{q}^{\tilde{L}_0 - \frac{\tilde{c}}{24}} (-1)^{F + \tilde{F}} y^{J_0}) \ , \qquad q := e^{2\pi i \tau}, \ y := e^{2\pi i z}, \tag{2.4}$$

over the Ramond-Ramond (RR) g-twisted sector $\mathcal{H}_g^{RR}$. Here J_0 is the zero mode of a $\mathfrak{u}(1)$ current algebra contained in the left $\mathcal{N} = 2$ superconformal algebra (normalized so that the charges are integral) and $(-1)^{F + \tilde{F}}$ is the total fermion number. This trace is computed by a path-integral with the same g- and h-twisted periodicity conditions for both bosonic and fermionic fields, with the insertion of an operator y^{J_0}. Although these conformal field theories are not chiral, the twisted twining partition functions are holomorphic with respect to both τ and z and are expected to be Jacobi forms of weight 0 and index $\frac{c}{6}$ under the same subgroup $\Gamma_{g,h} \subset SL(2, \mathbb{Z})$ as above.

2.2 Modular properties and the α-twist

We mentioned above that for a holomorphic CFT $\mathcal{C}$ with automorphism group G, the twisted sectors $\mathcal{H}_g$, carry projective representations ρ_g of the centralizer $C_G(g)$. This implies that ρ_g satisfies

$$\rho_g(h)\rho_g(h') = c_g(h,h')\rho_g(hh'), \qquad h,h' \in C_G(g) \ , \tag{2.5}$$

where $c_g(h,h')$ is a phase, or, more precisely, a 2-cocycle representing a class $[c_g]$ in $H^2(C_G(g),U(1))$. As a consequence, the twisted twining partition function $Z_{g,h}$ is not an honest class function on G, but rather satisfies

$$Z_{g,h}(\tau) = \frac{c_g(h,k)}{c_g(k,k^{-1}hk)} Z_{g,k^{-1}hk}(\tau), \qquad k \in C_G(g), \tag{2.6}$$

which puts strong constraints on $Z_{g,h}$ (see appendix B and [18]). The same formula also applies to the twisted twining elliptic genera $\phi_{g,h}$.

The role of the cohomology group $H^2(C_G(g),U(1))$ can in fact be derived from a more fundamental property of $\mathcal{C}$, namely that consistent orbifolds of a holomorphic CFT are classified by the third cohomology $H^3(G,U(1))$. Concretely, this implies there exists a $U(1)$-valued 3-cocycle $\alpha(g,h,k)$, with $(g,h,k) \in G\times G\times G$, representing a class $[\alpha]$ in $H^3(G,U(1))$, which in turn determines the class $[c_g] \in H^2(C_G(g),U(1))$ via the formula [47, 48]

$$c_g(x,y) := \frac{\alpha(g,x,y)\alpha(x,y,(xy)^{-1}g(xy))}{\alpha(x,x^{-1}gx,y)} \ , \qquad g,x,y \in G \ . \tag{2.7}$$

Thus, α characterizes the projective representations ρ_g. The 3-cocycle α is normalized, i.e. $\alpha(g,h,k) = 1$ whenever at least one of its arguments g,h,k is the identity. The cohomology group $H^3(G,U(1))$ also controls the modular properties of the twisted twining partition functions $Z_{g,h}$, in the sense that under a modular transformation $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in SL(2,\mathbb{Z})$ they satisfy

$$\epsilon_{g,h}\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) Z_{g^d h^{-c}, g^{-b}h^a}\left(\frac{a\tau+b}{c\tau+d}\right) = Z_{g,h}(\tau), \tag{2.8}$$

where $\epsilon_{g,h}\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ is a $U(1)$-valued multiplier system[‡] for $SL(2,\mathbb{Z})$ that can be explicitly constructed from the 3-cocycle $\alpha(g,h,k)$ (see appendix B). Furthermore, (2.6) generalizes to

$$Z_{g,h}(\tau) = \frac{c_g(h,k)}{c_g(k,k^{-1}hk)} Z_{k^{-1}gk,k^{-1}hk}(\tau), \qquad k \in G, \tag{2.9}$$

where conjugation is taken within the whole group G. In [18, 17] these properties were used to constrain and determine all of the twisted twining elliptic genera $\phi_{g,h}$ for g,h a commuting pair in M_{24}.

[‡] This was called $\chi_{g,h}$ in [18].

Given the rather complicated modular action in (2.8) it turns out to be very convenient to introduce a certain slash operator which combines the action of $SL(2,\mathbb{Z})$ on $(\tau; g, h)$ with the conjugation by G in (2.9). In addition this must incorporate the multiplier induced by the cocycle twist. To define the slash operator we must first define the following twisted modular action on commuting pairs of elements (g,h) in G:

$$(\gamma, k) : (g,h) \mapsto (kgk^{-1}, khk^{-1})\gamma^{-1} = (kg^d h^{-c} k^{-1}, kg^{-b} h^a k^{-1}) \ . \tag{2.10}$$

Here, $(\gamma, k) = \left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right), k \right) \in SL(2,\mathbb{Z}) \times G$.

When acting on twisted twining partition functions $Z_{g,h}$ the α-twisted equivariant slash operator is defined as follows

$$Z_{g,h}(\tau)|_\alpha(\gamma,k) := \epsilon_{g,h}(\gamma,k) Z_{(\gamma,k)\cdot(g,h)}(\gamma \cdot \tau) \ , \qquad (\gamma,k) \in SL(2,\mathbb{Z}) \times G \ , \tag{2.11}$$

where $\epsilon_{g,h}(\gamma,k) \in U(1)$ is a phase which depends on the choice of 3-cocycle α representing a class $[\alpha] \in H^3(G, U(1))$ (see appendix B.2 for more details). When $k = e$ this phase reduces to the phase $\epsilon_{g,h}(\gamma, e) = \epsilon_{g,h}(\gamma)$ in (2.8), while for $\gamma = 1$ we have that $\epsilon_{g,h}(1,k)$ is identified with the phase in (2.9).[§]

The class $[\alpha] \in H^3(G, U(1))$ is part of the defining data of a holomorphic CFT $\mathcal{C}$ with finite automorphism group G. In general the twisted twining partition functions $Z_{g,h}(\tau)$ associated with $\mathcal{C}$ are invariant under the α-twisted slash operator

$$Z_{g,h}(\tau)|_\alpha(\gamma,k) = Z_{g,h}(\tau). \tag{2.12}$$

Different choices of representative α in the class $[\alpha]$ are related to each other by rescalings of $Z_{g,h}(\tau)$ by overall (g,h)-dependent phases. More precisely, if two normalized 3-cocycles α and α' differ by a 3-coboundary $\partial\beta$, i.e.

$$\alpha(g_1,g_2,g_3) = \alpha'(g_1,g_2,g_3) \, \frac{\beta(g_1, g_2 g_3)\beta(g_2,g_3)}{\beta(g_1 g_2, g_3)\beta(g_1,g_2)} \ , \tag{2.13}$$

for some $\beta : G \times G \to U(1)$, with $\beta(e,g) = \beta(g,e) = 1$ for all $g \in G$, then the corresponding twisted twining partition functions $Z_{g,h}$ and $Z'_{g,h}$ are related by

$$Z'_{g,h}(\tau,z) = \frac{\beta(g,h)}{\beta(h,g)} Z_{g,h}(\tau,z) \ . \tag{2.14}$$

We also define the α-twisted slash operators on the twisted twining genera $\phi_{g,h}$ by the natural generalization of (2.11) to Jacobi forms of weight 0 and index 1:

$$\phi_{g,h}(\tau,z)|_\alpha(\gamma,k) := \epsilon_{g,h}(\gamma,k) e^{-\frac{2\pi i m c z^2}{c\tau+d}} \phi_{k\,g^d h^{-c}\,k^{-1},\,k\,g^{-b}h^a\,k^{-1}}(\gamma\cdot\tau, \frac{z}{c\tau+d}). \tag{2.15}$$

[§] This phase was called $\xi_{g,h}$ in [18].

2.3 (Generalized) Mathieu moonshine

The elliptic genus of K3 is defined in physics as a refined partition function

$$\phi_{K3}(\tau,z)=\mathrm{Tr}_{\mathcal{H}}(q^{L_0-\frac{c}{24}}\bar{q}^{\tilde{L}_0-\frac{\tilde{c}}{24}}(-1)^{F+\tilde{F}}y^{J_0})\ ,\qquad q:=e^{2\pi i\tau},\ y:=e^{2\pi iz},\tag{2.16}$$

of certain conformal field theories with $\mathcal{N}=(4,4)$ superconformal algebra at central charge $c=6$, namely non-linear sigma models with target space K3. The only states contributing to this partition function are the Ramond-Ramond right-moving ground states, i.e. with $\bar{L}_0-\frac{\bar{c}}{24}=0$.

In [1], Eguchi, Ooguri and Tachikawa conjectured a relationship (subsequently dubbed Mathieu moonshine) between the elliptic genus of K3 and the largest Mathieu group M_{24}. It is useful to formulate the Mathieu moonshine conjecture (now a theorem, thanks to [12]) in terms of an abstract representation $\mathcal{H}$ of the (holomorphic) $\mathcal{N}=4$ superconformal algebra at central charge $c=6$, which is also a module for the Mathieu group M_{24}. Heuristically, $\mathcal{H}$ can be interpreted as the spectrum of R-R right-moving ground states in a generic K3 model, although it is known that no such model has M_{24} as its symmetry group [8, 11]. The module $\mathcal{H}$ also admits a $\mathbb{Z}_2$-grading by a 'right-moving fermion number' $(-1)^{\bar{F}}$, which is preserved both by the $\mathcal{N}=4$ superconformal algebra and by the action of M_{24}. The precise statement of the conjecture is as follows.

Theorem (Mathieu moonshine). *There exists unitary Ramond representation $\mathcal{H}$ of the $\mathcal{N}=4$ superconformal algebra at central charge $c=6$, $\mathbb{Z}_2$-graded by the 'right-moving fermion number' $(-1)^{\bar{F}}$, that carries a non-trivial action of the Mathieu M_{24} commuting with both the $\mathcal{N}=4$ algebra and the $\mathbb{Z}_2$-grading. Furthermore, for each $g\in M_{24}$, the twining genus*

$$\phi_g(\tau,z):=\mathrm{Tr}_{\mathcal{H}}(g\,q^{L_0-\frac{c}{24}}y^{J_0^3}(-1)^{F+\bar{F}})\ ,\qquad g\in M_{24},\tag{2.17}$$

($q:=e^{2\pi i\tau}$, $y:=e^{2\pi iz}$) is a weak Jacobi form of weight 0 and index 1 (possibly with multiplier) under $\Gamma_0(N)$, where N is the order of g. In particular, ϕ_e is the elliptic genus of K3.

A list of Jacobi forms ϕ_g, $g\in M_{24}$, satisfying the expected modular properties was proposed in [2]–[5] (see table 1 in appendix F for the explicit expressions) and the existence of the corresponding module $\mathcal{H}$ was proved in [12]. The module $\mathcal{H}$ is not uniquely determined by this description, since one can always adjoin to $\mathcal{H}$ a pair of isomorphic representations of the $\mathcal{N}=4$ algebra and M_{24} with opposite right-moving fermion number, so that they do not contribute to any twisted twining genus. However, there is a 'minimal' module that contains no such pair of representations and this is uniquely determined by the twining genera ϕ_g. Remarkably, in this minimal module, the only $\mathcal{N}=4$ representations with negative $(-1)^{\bar{F}}$

are the BPS ones. It is believed that the module $\mathcal{H}$ and the Jacobi forms ϕ_g of [2]–[5] are the unique solutions of the Mathieu moonshine conjecture, although this has not been proved yet.

In the same spirit, generalized Mathieu moonshine can be expressed in terms of the existence of twisted modules $\mathcal{H}_g$ and twisted twining genera $\phi_{g,h}$.

Conjecture (Generalized Mathieu moonshine). *For each $g \in M_{24}$, there exists a $\mathbb{Z}_2$-graded unitary Ramond representation $\mathcal{H}_g$ of the $\mathcal{N} = 4$ algebra at central charge $c = 6$, that carries a projective representation $\rho_g \colon C_{M_{24}}(g) \to GL(\mathcal{H}_g)$ of the centralizer of g in M_{24}, with group law*

$$\rho_g(h)\rho_g(k) = c_g(h,k)\rho_g(hk) \ , \qquad h,k \in C_{M_{24}}(g) \ , \tag{2.18}$$

where $c_g(h,k)$ is defined by (2.7) in terms of a normalized 3-cocyle α representing a primitive class in $H^3(M_{24}, U(1)) \cong \mathbb{Z}_{12}$. The operators $\rho_g(h)$ commute with the $\mathcal{N} = 4$ algebra and with the right-moving fermion number $(-1)^{\bar{F}}$, and we have

$$\rho_g(g) = e^{2\pi i (L_0 - \frac{c}{24})} \ . \tag{2.19}$$

Furthermore, for each pair of commuting elements $g, h \in M_{24}$, the function

$$\phi_{g,h}(\tau, z) := \mathrm{Tr}_{\mathcal{H}_g}(\rho_g(h) q^{L_0 - \frac{c}{24}} y^{J_0^3} (-1)^{F+\bar{F}}) \ , \tag{2.20}$$

satisfies the following properties:

(A) For $g = e$, where e is the identity element of M_{24}, the functions $\phi_{e,h}$ coincide with the twining genera ϕ_h constructed in [2]–[5]. In particular, $\phi_{e,e}$ is the elliptic genus of $K3$.

(B) Elliptic properties:

$$\phi_{g,h}(\tau, z + \ell\tau + \ell') = e^{-2\pi i (\ell^2 \tau + 2\ell z)} \phi_{g,h}(\tau, z) \qquad \ell, \ell' \in \mathbb{Z} \ . \tag{2.21}$$

(C) Invariance under the α-twisted slash operator (see eq.(2.15))

$$\phi_{g,h}(\tau, z)|_\alpha(\gamma, k) = \phi_{g,h}(\tau, z) \ , \qquad (\gamma, k) \in SL(2,\mathbb{Z}) \times M_{24} \ . \tag{2.22}$$

In particular, each $\phi_{g,h}$ is an even weak Jacobi form of weight 0 and index 1 under a group $\Gamma_{g,h} \subseteq PSL(2,\mathbb{Z})$, with multiplier $\chi_{g,h}$ depending on α.

These conditions were presented in [17] where all the functions $\phi_{g,h}$ were explicitly found, and strong numerical evidence was given that they decompose into projective characters of

$C_{M_{24}}(g)$ with respect to the $\mathcal{N}=4$ algebra. It was also proven in [17] that there exists a unique cohomology class $[\alpha] \in H^3(M_{24}, U(1))$ for which (A)-(C) can be satisfied and that, for each choice of a normalized representative α, the functions $\phi_{g,h}$ satisfying (A)-(C) are uniquely determined. If two normalized 3-cocycles α and α' differ by a 3-coboundary $\partial\beta$, as in (2.13), then the corresponding twisted twining genera $\phi_{g,h}$ and $\phi'_{g,h}$ are related as in (2.14)

$$\phi'_{g,h}(\tau, z) = \frac{\beta(g,h)}{\beta(h,g)}\phi_{g,h}(\tau, z) \ . \tag{2.23}$$

The explicit expressions for $\phi_{g,h}$ derived in [17] are collected in table 2 in appendix F. Although the existence of the modules $\mathcal{H}_g$ matching (2.20) is not proven yet for g different from the identity, strong evidence in this direction has been given [17]. In the rest of the paper we will assume that the conjecture holds. As in the ordinary moonshine case, one can choose minimal modules $\mathcal{H}_g$, such that the only states with negative $(-1)^{\bar{F}}$ are contained in irreducible BPS representations of $\mathcal{N}=4$. In fact, the results of [17] suggest that such states only appear in the untwisted sector $\mathcal{H}_{g=e}$.

2.4 Geometric perspective

Before we close this section we shall offer a more geometric perspective on the twisted twining genera which will be useful later on. In [55], Ganter showed that the natural home for the Norton series $f(g,h;\tau)$ [19] is the equivariant elliptic cohomology developed in [56]. We shall here give a short review of Ganter's perspective, adapted to the case of Jacobi forms on $\mathbb{H}_+ \times \mathbb{C}$, as opposed to modular forms on $\mathbb{H}_+$.

Let $\mathcal{P}$ be the set of pairs of commuting elements of M_{24}

$$\mathcal{P} := \{(g,h) \in M_{24} \mid gh = hg\} \ , \tag{2.24}$$

and $\bar{\mathcal{P}}$ the set of conjugacy classes of such pairs, i.e. the quotient of $\mathcal{P}$ with respect to $(g,h) \sim (k^{-1}gk, k^{-1}hk)$, for any $k \in M_{24}$. The set $\bar{\mathcal{P}}$ can be identified with the set of isomorphism classes of principal M_{24}-bundles over the elliptic curve $E_\tau = \mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$. One can then consider the associated moduli space[¶] of principal M_{24}-bundles on E_τ

$$\mathcal{M} = \mathcal{P} \times (\mathbb{H}_+ \times \mathbb{C}) \,/\, M_{24} \times (SL(2,\mathbb{Z}) \ltimes \mathbb{Z}^2) \ . \tag{2.25}$$

For each pair $(g,h) \in \mathcal{P}$ there exists an $(SL(2,\mathbb{Z}) \times \mathbb{Z}^2)$-equivariant line bundle

$$\begin{tikzcd} \mathcal{L}^{\alpha}_{g,h} \arrow[d] \\ \mathcal{M} \end{tikzcd} \tag{2.26}$$

[¶] This should really be a moduli *stack* but we ignore this technical point; see [55] for a more precise description.

which is twisted by the 3-cocycle $\alpha \in H^3(M_{24}, U(1))$ [55]. Thus, for fixed (g,h) we can think of the twisted twining genus $\phi_{g,h}$ as a section of $\mathcal{L}^{\alpha}_{g,h}$. The twist by α in the bundle $\mathcal{L}^{\alpha}_{g,h}$ accounts for the α-dependent multiplier phases occurring in the modular transformations of the twisted twining genera. In the cases when α describes a trivial class in $H^3(M_{24}, U(1))$, the bundle $\mathcal{L}^{\alpha}_{g,h}$ is canonically trivialized and the associated $\phi_{g,h}$ exhibits no multiplier phase. The interpretation of the twisted twining genera as sections of $\mathcal{L}^{\alpha}_{g,h}$ will in particular play a role in section 3.3 when we discuss twisted equivariant Hecke operators.

3 Second quantization of Mathieu moonshine

In this section we construct the second quantized twisted twining genera $\Psi_{g,h}$ as the exponentiated generating function of certain twisted equivariant Hecke operators acting on $\phi_{g,h}$. The forms $\Psi_{g,h}$ have a natural interpretation in terms of generating functions of the twisted twining genera in the symmetric orbifold CFTs and we therefore begin by reviewing some relevant facts about symmetric orbifold theory. Having defined the functions $\Psi_{g,h}$ we also show that they have infinite product expansions.

3.1 Symmetric Orbifolds and Hecke Operators

Given a holomorphic CFT, or self-dual Vertex Operator Algebra (VOA), $\mathcal{C}$ with a group $G \subseteq \mathrm{Aut}(\mathcal{C})$ of automorphisms, let us consider the theory $\mathcal{C}^{\otimes L}$ obtained by taking the tensor product of L copies of $\mathcal{C}$, for some $L \geq 1$. The group of automorphisms of $\mathcal{C}^{\otimes L}$ contains the direct product of the symmetric group S_L and the group G acting diagonally on all copies of $\mathcal{C}$.[‖] The twisted twining partition function $Z^{\mathcal{C}^{\otimes L}}_{g,h}$ of the theory $\mathcal{C}^{\otimes L}$, associated to a pair of commuting elements (g,h) of the diagonal group G, is simply the L:th-power of the twisted twining partition function $Z_{g,h}$ in the original theory $\mathcal{C}$:

$$Z^{\mathcal{C}^{\otimes L}}_{g,h}(\tau) = Z_{g,h}(\tau)^L \ . \tag{3.1}$$

In other words, the path integral on the torus with boundary conditions g, h in the product theory $\mathcal{C}^{\otimes L}$ is simply the product of L copies of the path integral in the original theory with the same boundary conditions. In particular, if $Z_{g,h}$ is invariant under the α-twisted slash operator for some 3-cocycle α, then $Z^{\mathcal{C}^{\otimes L}}_{g,h}$ is invariant under the α^L-twisted slash operator.

[‖] Actually, $\mathrm{Aut}(\mathcal{C}^{\otimes L})$ contains the wreath product $S_L \wr G := S_L \ltimes (G \times \ldots \times G)$, i.e. the semidirect product of S_L and the direct product $G^L := G \times \ldots \times G$ of L copies of G, with S_L acting on G^L by the obvious permutation.

The L:th symmetric orbifold $S^L\mathcal{C}$ of $\mathcal{C}$ is defined as the orbifold of $\mathcal{C}^{\otimes L}$ by S_L, and its group of automorphisms contains the diagonal G. Therefore, one can define the twisted twining partition functions

$$Z^{(L)}_{g,h}(\tau) \equiv Z^{S^L\mathcal{C}}_{g,h}(\tau) \tag{3.2}$$

in each symmetric product. These partition functions can be computed using the usual formulae for orbifold CFTs.

The orbifold formula for $Z^{(L)}_{g,h}$ can be nicely understood in the geometric setting of section 2.4. Recall that the twisted twining partition function $Z_{g,h}(\tau)$ of a VOA $\mathcal{C}$ with $\mathrm{Aut}(\mathcal{C}) = G$ is a section of a (α-twisted) line bundle $\mathcal{L}^\alpha_{g,h}$ over the moduli space $\mathcal{M}_G$ of G-bundles over the elliptic curve $E = \mathbb{C}/\Lambda$ with $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$. In this picture g and h correspond to the monodromies around the cycles -1 and τ of the elliptic curve.

As shown in [23], $Z^{(L)}_{g,h}$ can be expressed as a sum of contributions from all isomorphism classes of unramified L-fold coverings $\Upsilon : E' \to E$, namely

$$Z^{(L)}_{g,h}(\tau) \sim \sum_{\substack{\text{isomorphism classes of}\\ L\text{-fold coverings}\\ \Upsilon: E'\to E}} \Upsilon^* Z_{g,h}(\tau) \ , \tag{3.3}$$

up to a suitable normalization. Here, $\Upsilon^* Z_{g,h}$ is simply the partition function associated with the pull-back $\Upsilon^*\mathcal{L}^\alpha_{g,h} \to E'$ of the line bundle $\mathcal{L}^\alpha_{g,h}$.

Up to isomorphisms, we can consider coverings that preserve the base-point of the elliptic curve. Any *connected* unramified base-point preserving L-fold cover of $E = \mathbb{C}/\Lambda$ is given by $E' \cong \mathbb{C}/\Lambda' \to E$, $z \mapsto z$, where Λ' is a sublattice of index L in Λ. In turn, the sublattices of index L are given by $\Lambda' = M\Lambda$ for any M in the set

$$\mathrm{Mat}_L(\mathbb{Z}) := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{Z}, ad - bc = L\} \tag{3.4}$$

of integral 2×2 matrices M of determinant L. In particular, when $L = 1$, the covers $\Upsilon : E' \to E$ are actually isomorphisms and are classified by $\mathrm{Mat}_1(\mathbb{Z}) \equiv SL(2,\mathbb{Z})$. Let Υ_γ denote the isomorphism associated with a specific choice of $\gamma \in SL(2,\mathbb{Z})$. The pull-back $\Upsilon^*_\gamma Z_{g,h}$ then coincides with the slash operator

$$\Upsilon^*_\gamma Z_{g,h}(\tau) = Z_{g,h}(\tau)|_\alpha(\gamma, e) \ , \qquad \gamma \in SL(2,\mathbb{Z}) \ . \tag{3.5}$$

For generic $L > 0$, the isomorphism classes of coverings are in one to one correspondence with the cosets

$$SL(2,\mathbb{Z})\backslash \mathrm{Mat}_L(\mathbb{Z}) \ , \tag{3.6}$$

and a set of coset representatives is given by

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} , \qquad ad = L,\ a,d > 0,\ 0 \le b < d\ . \tag{3.7}$$

Therefore, the contribution to the partition function $Z^{(L)}_{g,h}(\tau)$ from the *connected* L-fold coverings is given by

$$\mathcal{T}^{\alpha}_{L} Z_{g,h}(\tau) := \frac{1}{L} \sum_{M\in SL_2(\mathbb{Z})\backslash \mathrm{Mat}_L(\mathbb{Z})} \Upsilon^{*}_{M} Z_{g,h}(\tau) \equiv \frac{1}{L} \sum_{M\in SL_2(\mathbb{Z})\backslash \mathrm{Mat}_L(\mathbb{Z})} Z_{g,h}(\tau)|_{\alpha}(M,e)\ , \tag{3.8}$$

where $|_{\alpha}(M,e)$ is a suitable generalization of the slash operator to $\mathrm{Mat}_L(\mathbb{Z})$, that will be discussed in the following sections and in appendix C.

The full twisted twining partition function $Z^{(L)}_{g,h}$ is given by the sum over connected and disconnected L-coverings. The result is most easily expressed in terms of a generating function, which is given by the plethystic exponential of the connected contribution

$$\sum_{L=0}^{\infty} p^L Z^{(L)}_{g,h}(\tau) = \exp\Big(\sum_{L=1}^{\infty} p^L \mathcal{T}^{\alpha}_{L} Z_{g,h}(\tau)\Big)\ . \tag{3.9}$$

Explicit formulae for the (α-twisted) equivariant Hecke operators $\mathcal{T}^{\alpha}_{L}$ defined by (3.8) will be discussed in more detail in section 3.2 and appendix C.

3.2 Twisted equivariant Hecke operators

The exponential lifts $\Psi_{g,h}$ of the twisted twining genera $\phi_{g,h}$ are the analogue of the generating functions (3.9) for superconformal field theories. In particular, they can be defined in terms of some set of Hecke operators that are compatible with the properties (A)-(C) in section 2.3. This means that the Hecke action should be compatible, i.e. equivariant, with respect to the M_{24}-action on the pair (g,h), and it should incorporate the α-twist in the modular transformation. Equivariant versions of Hecke operators acting on modular forms was proposed by Ganter [55] in the context of generalized Monstrous moonshine, and we shall see that this result also applies here after some minor modifications.

3.2.1 Generalities on Hecke operators

Hecke operators are linear operators acting on modular forms. Specifically, for each $L \ge 1$ there is an operator T_L that acts on a weight k modular form $f(\tau)$ and produces another modular form $T_L f(\tau)$ of the same weight. On modular functions, i.e. weight $k=0$ modular forms, the Hecke operator is defined by

$$T_L f(\tau) := \frac{1}{L} \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \backslash \mathrm{Mat}_L(\mathbb{Z})} f\left(\frac{a\tau+b}{c\tau+d}\right) . \tag{3.10}$$

These operators have a natural generalization to Jacobi forms given by [57] (restricting to the weight 0, index m case)

$$T_L \phi(\tau, z) := \frac{1}{L} \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \backslash \mathrm{Mat}_L(\mathbb{Z})} e^{\frac{-mLcz^2}{c\tau+d}} \phi\left(\frac{a\tau+b}{c\tau+d}, \frac{Lz}{c\tau+d}\right) . \tag{3.11}$$

In fact, on Jacobi forms there exists an additional Hecke operator U_L with the simple action

$$U_L \phi(\tau, z) := \phi(\tau, Lz). \tag{3.12}$$

These operators map Jacobi forms of weight 0 and index m to Jacobi forms of weight 0 and index Lm and $L^2 m$, respectively. They also form a Hecke algebra, with the relations

$$T_m \cdot T_n = T_{mn} \ , \qquad \text{for } \gcd(m,n) = 1 \ , \tag{3.13}$$

$$U_m \cdot U_n = U_{mn} \ , \tag{3.14}$$

$$T_m \cdot U_n = U_n \cdot T_m \ , \tag{3.15}$$

$$T_p \cdot T_{p^m} = T_{p^{m+1}} + \frac{1}{p} T_{p^{m-1}} \cdot U_p \ , \qquad \text{for } p \text{ prime} \ . \tag{3.16}$$

It is easy to see that a set of representatives for the quotient $SL(2,\mathbb{Z}) \backslash \mathrm{Mat}_L(\mathbb{Z})$ can always be chosen of the form (3.7). This allows us to write the action of T_L on $\phi(\tau, z)$ as follows

$$T_L \phi(\tau, z) = \frac{1}{L} \sum_{a,d>0, ad=L} \sum_{b=0}^{d-1} \phi\left(\frac{a\tau+b}{d}, az\right) . \tag{3.17}$$

3.2.2 Twisted equivariant Hecke operators: main definition

The standard Hecke-operators T_L map weak Jacobi forms of weight 0, index 1 to weak Jacobi forms of weight 0, index L. In the generalized setting the analogous statement implies that we should have equivariant Hecke operators $\mathcal{T}_L$ that map sections of $\mathcal{L}^{\alpha}_{g,h}$ to sections of the product bundle $(\mathcal{L}^{\alpha}_{g,h})^{\otimes L}$, where $\mathcal{L}^{\alpha}_{g,h} \to \mathcal{M}$ is the line bundle discussed in section 2.4. In fact, because of the cocycle-dependent multiplier phases arising in the modular transformations of the $\phi_{g,h}$ we need to introduce a certain α-twisted version $\mathcal{T}^{\alpha}_L$ of the equivariant Hecke operator. More explicitly, if the sections of $\mathcal{L}^{\alpha}_{g,h}$ correspond to holomorphic functions $\phi_{g,h}$ on the covering space $\mathcal{P} \times \mathbb{H}_+ \times \mathbb{C}$ of $\mathcal{M}$, with an automorphy factor determined by a 3-cocycle α, then $\mathcal{T}^{\alpha}_L \phi_{g,h}$ should correspond to holomorphic functions whose automorphy factor derives from the L:th power α^L of the 3-cocycle.

Since the detailed analysis is rather technical, we will postpone it to appendix C and just quote the explicit formula for the action of the Hecke operators on a twisted twining character $\phi_{g,h}$

$$\mathcal{T}^{\alpha}_L \phi_{g,h}(\tau, z) := \frac{1}{L} \sum_{\substack{a,d>0,\\ ad=L}} \sum_{b=0}^{d-1} \epsilon_{g,h} \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix} \right) \phi_{g^d, g^{-b}h^a} \Big(\frac{a\tau + b}{d}, az \Big) \ , \tag{3.18}$$

$$\mathcal{U}^{\alpha}_L \phi_{g,h}(\tau, z) := \epsilon_{g,h} \left(\begin{smallmatrix} L & 0 \\ 0 & L \end{smallmatrix} \right) \phi_{g^L, h^L}(\tau, Lz) \ , \tag{3.19}$$

where

$$\epsilon_{g,h} \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix} \right) := \frac{\prod_{i=1}^{a-1} c_g(h, h^i)^d}{\prod_{j=1}^{d-1} c_{g^{-b}h^a}(g, g^j) \prod_{k=1}^{b} c_g(g, g^{-k}h^a)^d} \ , \tag{3.20}$$

and $c_g(x,y)$, $g, x, y \in G$, depends on the cocycle α that determines the automorphy factor of $\phi_{g,h}$ via (2.7).

As discussed in appendix C, these Hecke operators satisfy the Hecke algebra (3.13)–(3.16). The explicit expressions for $\mathcal{T}^{\alpha}_L \phi_{g,h}$, for $L = 1, 2, 3, 4$, in the case of Mathieu moonshine are collected in appendix F.

Notice that the phases $\epsilon_{g,h} \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right)$ are trivial whenever g is the identity or when the restriction of the cocycle α to the group $\langle g, h \rangle$ is trivial. An example where non-trivial phases appear is given by $\mathcal{T}^{\alpha}_2 \phi_{g,g}$, where g is in class $2B$ of M_{24}. These phases can be understood by first considering

$$\mathcal{T}^{\alpha}_2 \phi_{e,g}(\tau, z) = \frac{1}{2} \Big(\phi_{e,e}(2\tau, 2z) + \phi_{e,g}(\frac{\tau}{2}, z) + \phi_{e,g}(\frac{\tau+1}{2}, z) \Big) \ , \tag{3.21}$$

and then imposing the relations

$$\mathcal{T}^{\alpha}_2 \phi_{g,e}(\tau, z) = e^{\frac{-4\pi i z^2}{\tau}} \mathcal{T}^{\alpha}_2 \phi_{e,g}(-\frac{1}{\tau}, \frac{z}{\tau}) \ , \qquad \mathcal{T}^{\alpha}_2 \phi_{g,g}(\tau, z) = \mathcal{T}^{\alpha}_2 \phi_{g,e}(\tau + 1, z) \ . \tag{3.22}$$

Using the modular properties of $\phi_{e,g}$, we obtain

$$\mathcal{T}^{\alpha}_2 \phi_{g,e}(\tau, z) = \frac{1}{2} \Big(\phi_{g,e}(2\tau, 2z) + \phi_{e,e}(\frac{\tau}{2}, z) - \phi_{e,g}(\frac{\tau+1}{2}, z) \Big) \ , \tag{3.23}$$

$$\mathcal{T}^{\alpha}_2 \phi_{g,g}(\tau, z) = \frac{1}{2} \Big(-\phi_{g,e}(2\tau, 2z) - \phi_{e,g}(\frac{\tau}{2}, z) + \phi_{e,e}(\frac{\tau+1}{2}, z) \Big) \ , \tag{3.24}$$

where the minus signs in the last two equations are due to the non-trivial multiplier system of the twining genus $\phi_{e,g}$.

3.2.3 Central extensions

Rather than considering the projective representation ρ_g of the centralizer $C_{M_{24}}(g)$, it is often useful to work with a linear representation of a central extension

$$1 \to U(1) \to C^{\alpha}_{M_{24}}(g) \to C_{M_{24}}(g) \to 1 \ . \tag{3.25}$$

The group $C^{\alpha}_{M_{24}}(g)$ can be constructed explicitly in terms of generators and relations. Let $q(x) := e^{2\pi i x}$, $x \in \mathbb{R}/\mathbb{Z}$, be the generator of the central $U(1)$ factor and consider one generator h_α for each $h \in C_{M_{24}}(g)$. Then, $C^{\alpha}_{M_{24}}(g)$ is generated by all such elements subject to the relations

$$h_\alpha k_\alpha = q\big(\mu_g(h,k)\big)\,(hk)_\alpha \ , \qquad q(x)h_\alpha = h_\alpha q(x) \tag{3.26}$$

for all $h, k \in C_{M_{24}}(g)$. Here, $\mu_g(h,k) \in \mathbb{R}/\mathbb{Z}$ is defined by

$$e^{2\pi i \mu_g(h,k)} = c_g(h,k) \ , \tag{3.27}$$

where $c_g(h,k)$ is the 2-cocycle determining the projective representation ρ_g as in (2.5). This definition provides also a canonical lift

$$h \mapsto h_\alpha \ , \tag{3.28}$$

from $C_{M_{24}}(g)$ to its central extension $C^{\alpha}_{M_{24}}(g)$. By construction, the g-twisted sector carries a genuine representation $\tilde{\rho}_g$ of $C^{\alpha}_{M_{24}}(g)$, with

$$\tilde{\rho}_g(q(x)) = e^{2\pi i x} \ , \qquad \tilde{\rho}_g(h_\alpha) = \rho_g(h) \ , \tag{3.29}$$

for all $h \in C_{M_{24}}(g)$, $x \in \mathbb{R}/\mathbb{Z}$. A less trivial observation is that also the g^r-twisted sectors, for all $r \in \mathbb{Z}_{\geq 0}$, carry a genuine representation $\tilde{\rho}_{g,r}$ of $C^{\alpha}_{M_{24}}(g)$, defined by[**]

$$\tilde{\rho}_{g,r}(q(x)) := e^{2\pi i r x} \ , \qquad \tilde{\rho}_{g,r}(h_\alpha) \equiv \tilde{\rho}_{g,r}(h) := \frac{\rho_{g^r}(h)}{f_{g,r}(h)} \ , \tag{3.30}$$

for all $h \in C_{M_{24}}(g) \subseteq C_{M_{24}}(g^r)$, $x \in \mathbb{R}/\mathbb{Z}$, $r \in \mathbb{Z}_{\geq 0}$, where

$$f_{g,r}(h) = \prod_{i=1}^{r-1} c_h(g, g^i) \ . \tag{3.31}$$

The fact that $\tilde{\rho}_{g,r}$ are well defined representations of $C^{\alpha}_{M_{24}}(g)$ is an immediate consequence of the identity

$$c_{g^r}(h,k) = \frac{f_{g,r}(h) f_{g,r}(k)}{f_{g,r}(hk)}\, c_g(h,k)^r \ , \qquad h, k \in C(g) \ , \tag{3.32}$$

[**] We denote by the same symbol $\tilde{\rho}_{g,r}$ also the projective representation of $C_{M_{24}}(g)$ on $\mathcal{H}_{g^r}$ with $\tilde{\rho}_{g,r}(h) := \rho_{g^r}(h)/f_{g,r}(h)$

which follows from the definition of c_g in terms of α and repeated applications of the cocycle condition for α. Notice that $f_{g,1}(h) = 1$ for all $h \in C_{M_{24}}(g)$, so that $\tilde{\rho}_g \equiv \tilde{\rho}_{g,1}$.

If α and α' are different cocycle representatives of the same cohomology class $[\alpha]$, then the central extensions are isomorphic $C^{\alpha}_{M_{24}}(g) \cong C^{\alpha'}_{M_{24}}(g)$, although the corresponding lifts h_α and $h_{\alpha'}$ are different. In particular, for special choices of the representative cocycle α, it is sufficient to consider a central extension of $C_{M_{24}}(g)$ by a finite subgroup of $U(1)$. With slight abuse of notation, we will denote also these finite central extensions by $C^{\alpha}_{M_{24}}(g)$. See appendix D.2 for more details on these special choices for α.

3.2.4 Twisted equivariant Hecke operators: alternative definition

The α-twisted equivariant Hecke operators are more easily defined in terms of the central extension $C^{\alpha}_{M_{24}}(g)$ described in section 3.2.3. Set

$$\phi_{g^r_\alpha, h_\alpha}(\tau, z) := \mathrm{Tr}_{\mathcal{H}_{g^r}}\left(\tilde{\rho}_{g,r}(h_\alpha) q^{L_0 - \frac{c}{24}} y^{J^3_0} (-1)^{F+\bar{F}}\right) . \tag{3.33}$$

so that

$$\begin{aligned}
\phi_{g^d_\alpha, g_\alpha^{-b} h^a_\alpha}(\tau, z) &= \mathrm{Tr}_{\mathcal{H}_{g^d}}\left(\tilde{\rho}_{g,d}(g)^{-b} \tilde{\rho}_{g,d}(h)^a q^{L_0 - \frac{c}{24}} y^{J^3_0} (-1)^{F+\bar{F}}\right) \\
&= \frac{\prod_{i=1}^{a-1} c_g(h, h^i)^d}{\prod_{k=1}^{b} c_g(g, g^{-k} h^a)^d} \mathrm{Tr}_{\mathcal{H}_{g^d}}\left(\tilde{\rho}_{g,d}(g^{-b} h^a) q^{L_0 - \frac{c}{24}} y^{J^3_0} (-1)^{F+\bar{F}}\right) \\
&= \frac{\prod_{i=1}^{a-1} c_g(h, h^i)^d}{\prod_{j=1}^{d-1} c_{g^{-b} h^a}(g, g^j) \prod_{k=1}^{b} c_g(g, g^{-k} h^a)^d} \mathrm{Tr}_{\mathcal{H}_{g^d}}\left(\rho_{g^d}(g^{-b} h^a) q^{L_0 - \frac{c}{24}} y^{J^3_0} (-1)^{F+\bar{F}}\right) \\
&= \epsilon_{g,h}\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \phi_{g^d, g^{-b} h^a}(\tau, z) .
\end{aligned} \tag{3.34}$$

Therefore, we can reinterpret the α-twisted Hecke operators $\mathcal{T}^{\alpha}_L$ acting on the twisted twining genera $\phi_{g,h}$ as (untwisted) equivariant operators acting on the twisted twining genera $\phi_{g_\alpha, h_\alpha}$

$$\mathcal{T}^{\alpha}_L \phi_{g,h}(\tau, z) = \frac{1}{L} \sum_{ad=L} \sum_{b=0}^{d-1} \phi_{g^d_\alpha, g_\alpha^{-b} h^a_\alpha}\left(\frac{a\tau + b}{d}, az\right) =: \mathcal{T}_L \phi_{g_\alpha, h_\alpha}(\tau, z) . \tag{3.35}$$

This form of the Hecke operators will turn out to be very useful for some of the calculations in sections 3.3 and 4.

3.3 Second-Quantized Twisted Twining Genera

3.3.1 Definition

For any Calabi-Yau manifold X, one can define its *second-quantized* elliptic genus Ψ_X as the exponentiated generating function of the orbifold elliptic genus $\phi_{S^L X}(\tau, z)$ of the symmetric products $S^L X$ [23]. In fact, ref. [23] gave three equivalent expressions for Ψ_X:

$$\Psi_X(\sigma, \tau, z) = \sum_{L=0}^{\infty} p^L \phi_{S^L X}(\tau, z) = \exp\Big[\sum_{L=1}^{\infty} p^L (T_L \phi_X)(\tau, z)\Big] = \prod_{\substack{n>0, m\geq 0,\\ \ell \in \mathbb{Z}}} \big(1 - p^n q^m y^\ell\big)^{-c(mn,\ell)}, \tag{3.36}$$

where we have set

$$q = e^{2\pi i \tau}, \qquad y = e^{2\pi i z}, \qquad p = e^{2\pi i \rho}, \tag{3.37}$$

and $c(mn, \ell)$ are the Fourier coefficients of the elliptic genus of X:

$$\phi_X(\tau, z) = \sum_{m\geq 0, \ell\in\mathbb{Z}} c(m, \ell) q^m y^\ell. \tag{3.38}$$

Here, T_L are standard Hecke operators defined in (3.17).

It was shown in [50] that $\Phi \equiv A_X/\Psi_X$, where A_X is a simple correction factor ("Hodge anomaly") that is determined by the Hodge numbers of X, transforms as a Siegel modular form for some congruence subgroup of $Sp(4; \mathbb{Z})$.

Taking the formula (3.36) as a starting point, we now wish to define the *second quantized twisted twining genera* as follows

$$\Psi_{g,h}(\sigma, \tau, z) := \exp\Big[\sum_{L=1}^{\infty} p^L (\mathcal{T}_L^\alpha \phi_{g,h})(\tau, z)\Big], \tag{3.39}$$

where $\mathcal{T}_L^\alpha$ is the twisted equivariant Hecke operator defined in (3.18). We will later show that after including a correction factor $A_{g,h}$ the functions $\Psi_{g,h}$ transform as Siegel modular forms for some discrete subgroup $\Gamma^{(2)}_{g,h} \subset Sp(4; \mathbb{R})$ that contains the invariance group $\Gamma_{g,h} \subset SL(2, \mathbb{Z})$ of the twisted twining genera $\phi_{g,h}$.

Notice that the definition (3.39) depends on the choice of a normalized 3-cocycle α. If α and α' differ by a 3-coboundary $\partial\beta$ as in(A.5), the corresponding twisted twining genera are related as

$$\mathcal{T}_L^{\alpha'} \phi'_{g,h}(\tau, z) = e^{2\pi i \nu_{g,h} L} \mathcal{T}_L^\alpha \phi_{g,h}(\tau, z) \ , \tag{3.40}$$

where $\nu_{g,h} \in \mathbb{R}/\mathbb{Z}$ is defined as

$$e^{2\pi i \nu_{g,h}} = \frac{\beta(g, h)}{\beta(h, g)} \ , \tag{3.41}$$

so that

$$\Psi'_{g,h}(\sigma,\tau,z) = \exp\Big(\sum_{L=1}^{\infty} p^L \mathcal{T}_L^{\alpha'} \phi'_{g,h}(\tau,z)\Big) \tag{3.42}$$

$$= \exp\Big(\sum_{L=1}^{\infty} (e^{2\pi i \nu_{g,h}} p)^L\, \mathcal{T}_L^{\alpha} \phi_{g,h}(\tau,z)\Big) = \Psi_{g,h}(\sigma+\nu_{g,h},\tau,z)\ . \tag{3.43}$$

Therefore, a different choice for the cocycle representative α simply amounts to a redefinition $\sigma \to \sigma + \nu_{g,h}$ of the variable σ.

3.3.2 Infinite product representation

We shall now derive infinite product representations for the second quantized twisted twining genera $\Psi_{g,h}$ in terms of the Fourier coefficients of $\phi_{g,h}$.

Let us consider a generic commuting pair of elements $g, h \in M_{24}$, with the cocycle α inducing a possibly non-trivial multiplier for $\phi_{g,h}$. Let $N = o(g)$ be the order of g and λ the length of the shortest cycle of g in the 24-dimensional permutation representation (see Table 1 in appendix F). Then, the lift g_α of g to the central extension $C^{\alpha}_{M_{24}}(g)$, as defined in section 3.2.3, has order $N\lambda$ (see appendix D.2). For any $h \in C_{M_{24}}(g)$ the associated twisted twining genus $\phi_{g,h}$ has a Fourier expansion of the form

$$\phi_{g,h}(\tau,z) = \sum_{n=0}^{\infty}\sum_{\ell\in\mathbb{Z}} c_{g,h}\big(\frac{n}{N\lambda},\ell\big) q^{\frac{n}{N\lambda}} y^{\ell}\ , \tag{3.44}$$

where, in particular, $c_{g,h}\big(\frac{n}{N\lambda},\ell\big) = 0$ unless $n \equiv -1 \mod \lambda$. Let $M = o(h_\alpha)$ be the order of the lift h_α of h to the central extension $C^{\alpha}_{M_{24}}(g)$.[††] The logarithm of the second quantized twisted twining genus of $\phi_{g,h}$ is

---

[††] We assume that α is chosen in such a way that the order is finite.

$$\begin{aligned}\log\Psi_{g,h}(\sigma,\tau,z) &= \sum_{L=1}^{\infty} p^L \mathcal{T}_L^{\alpha}\phi_{g,h}(\tau,z)\\ &= \sum_{a,d=1}^{\infty}\frac{1}{ad}\sum_{b=0}^{d-1}\sum_{\ell\in\mathbb{Z}}\sum_{n=0}^{\infty}\epsilon_{g,h}\left(\begin{smallmatrix} a & b\\ 0 & d\end{smallmatrix}\right) c_{g^d,g^{-b}h^a}(\frac{n}{N\lambda},\ell)e^{\frac{2\pi i b n}{N\lambda d}}q^{\frac{an}{N\lambda d}}y^{a\ell}p^{ad}\\ &= \sum_{a,d=1}^{\infty}\frac{1}{ad}\sum_{\ell\in\mathbb{Z}}\sum_{n=0}^{\infty}\sum_{b=0}^{d-1} c_{g_\alpha^d,g_\alpha^{-b}h_\alpha^a}(\frac{n}{N\lambda},\ell)e^{\frac{2\pi i b n}{N\lambda d}}q^{\frac{an}{N\lambda d}}y^{a\ell}p^{ad}\\ &= \sum_{a,d=1}^{\infty}\frac{1}{a}\sum_{t,k=0}^{M-1}\frac{e^{\frac{2\pi i t(a-k)}{M}}}{M}\sum_{\ell\in\mathbb{Z}}\sum_{n=0}^{\infty}\frac{1}{d}\sum_{b=0}^{d-1} c_{g_\alpha^d,g_\alpha^{-b}h_\alpha^k}(\frac{n}{N\lambda},\ell)e^{\frac{2\pi i b n}{N\lambda d}}q^{\frac{an}{N\lambda d}}y^{a\ell}p^{ad}\\ &= \sum_{a,d=1}^{\infty}\frac{1}{a}\sum_{t,k=0}^{M-1}\frac{e^{\frac{2\pi i t(a-k)}{M}}}{M}\sum_{\ell\in\mathbb{Z}} F_{g,h}(a,d,k,\ell)y^{a\ell}p^{ad}\ ,\end{aligned} \tag{3.45}$$

where

$$F_{g,h}(a,d,k,\ell) := \sum_{n=0}^{\infty}\frac{1}{d}\sum_{b=0}^{d-1} e^{\frac{2\pi i b n}{N\lambda d}}\epsilon_{g,h}\left(\begin{smallmatrix} k & b\\ 0 & d\end{smallmatrix}\right) c_{g^d,g^{-b}h^k}(\frac{n}{N\lambda},\ell)q^{\frac{an}{dN\lambda}}\ . \tag{3.46}$$

In appendix D.1 we show that this sum can be rewritten as follows

$$F_{g,h}(a,d,k,\ell) = \sum_{m=0}^{\infty}\frac{1}{N\lambda}\sum_{b=0}^{\lambda N-1} e^{\frac{2\pi i b m}{N\lambda}}\epsilon_{g,h}\left(\begin{smallmatrix} k & b\\ 0 & d\end{smallmatrix}\right) c_{g^d,g^{-b}h^k}(\frac{md}{N\lambda},\ell)q^{\frac{am}{N\lambda}}\ . \tag{3.47}$$

By plugging this expression into (3.45) we obtain

$$\begin{aligned}\log\Psi_{g,h}(\sigma,\tau,z) =& \sum_{d=1}^{\infty}\sum_{m=0}^{\infty}\sum_{\ell\in\mathbb{Z}}\sum_{t=0}^{M-1}\hat{c}_{g,h}(d,m,\ell,t)\sum_{a=1}^{\infty}\frac{1}{a}\Big(e^{\frac{2\pi i t}{M}}q^{\frac{m}{N\lambda}}y^{\ell}p^{d}\Big)^a\\ &= -\sum_{d=1}^{\infty}\sum_{m=0}^{\infty}\sum_{\ell\in\mathbb{Z}}\sum_{t=0}^{M-1}\hat{c}_{g,h}(d,m,\ell,t)\log\big(1-e^{\frac{2\pi i t}{M}}q^{\frac{m}{N\lambda}}y^{\ell}p^{d}\big),\end{aligned} \tag{3.48}$$

where

$$\begin{aligned}\hat{c}_{g,h}(d,m,\ell,t) :=& \sum_{k=0}^{M-1}\sum_{b=0}^{\lambda N-1}\frac{e^{-\frac{2\pi i t k}{M}}}{M}\frac{e^{\frac{2\pi i b m}{\lambda N}}}{\lambda N}\epsilon_{g,h}\left(\begin{smallmatrix} k & b\\ 0 & d\end{smallmatrix}\right) c_{g^d,g^{-b}h^k}(\frac{md}{N\lambda},\ell)\\ =& \sum_{k=0}^{M-1}\sum_{b=0}^{\lambda N-1}\frac{e^{-\frac{2\pi i t k}{M}}}{M}\frac{e^{\frac{2\pi i b m}{\lambda N}}}{\lambda N}\mathrm{Tr}_{\mathcal{H}_{g^d}(\frac{md}{N\lambda},\ell)}\big(\tilde{\rho}_{g,d}(g)^{-b}\tilde{\rho}_{g,d}(h)^k(-1)^{F+\bar{F}}\big)\ .\end{aligned} \tag{3.49}$$

Eq.(3.49) implies that $\hat{c}_{g,h}(d,m,\ell,t)$ admit an interpretation as a $\mathbb{Z}_2$-graded dimension (with the grading given by $(-1)^{F+\bar{F}}$) of the simultaneous eigenspace for $\tilde{\rho}_{g,d}(g)$ and $\tilde{\rho}_{g,d}(h)$, restricted to $\mathcal{H}_{g^d}(\frac{md}{N\lambda},\ell)$, relative to the eigenvalues $e^{\frac{2\pi i m}{\lambda N}}$ and $e^{\frac{2\pi i t}{M}}$, respectively. In particular, $\hat{c}_{g,h}(d,m,\ell,t)$ is always an integer.

Thus, for the inverse of the second quantized twisted twining genus we obtain the infinite product expression

$$\frac{1}{\Psi_{g,h}(\sigma,\tau,z)} = \exp\left(-\sum_{L=1}^{\infty} p^L \mathcal{T}_L^{\alpha}\phi_{g,h}(\tau,z)\right) = \prod_{d=1}^{\infty}\prod_{m=0}^{\infty}\prod_{\ell\in\mathbb{Z}}\prod_{t=0}^{M-1}(1-e^{\frac{2\pi i t}{M}}q^{\frac{m}{N\lambda}}y^{\ell}p^{d})^{\hat{c}_{g,h}(d,m,\ell,t)} . \tag{3.50}$$

Note that (3.49) makes sense also for $d=0$, so that the infinite product in d and m can be symmetrized to obtain

$$\Phi_{g,h}(\sigma,\tau,z) := pq^{\frac{1}{N\lambda}}y \prod_{(d,m,\ell)>0}\prod_{t=0}^{M-1}(1-e^{\frac{2\pi i t}{M}}q^{\frac{m}{N\lambda}}y^{\ell}p^{d})^{\hat{c}_{g,h}(d,m,\ell,t)}, \tag{3.51}$$

where the first product runs over

$$d,m\in\mathbb{Z}_{\geq 0} \quad \text{and} \quad \begin{cases} \ell\in\mathbb{Z},\ \ell<0, & \text{if } m=0=d\ , \\ \ell\in\mathbb{Z}, & \text{otherwise.}\end{cases} \tag{3.52}$$

The prefactor $pq^{\frac{1}{N\lambda}}y$ in (3.51) has been chosen in such a way that $\Phi_{g,h}$ satisfies suitable automorphic properties, that will be described in section 5.2.

One can rewrite $\Phi_{g,h}$ in terms of the original $\Psi_{g,h}$ as

$$\Phi_{g,h}(\sigma,\tau,z) = \frac{p\psi_{g,h}(\tau,z)}{\Psi_{g,h}(\sigma,\tau,z)}, \tag{3.53}$$

where $\psi_{g,h}$ includes the $d=0$ factors in (3.51) and takes the form

$$\psi_{g,h}(\tau,z) = q^{\frac{1}{N\lambda}}y\prod_{t=0}^{M-1}\Big(\prod_{\ell<0}(1-e^{\frac{2\pi i t}{M}}y^{\ell})^{\hat{c}_{g,h}(0,0,\ell,t)}\Big)\Big(\prod_{\ell\in\mathbb{Z}}\prod_{m=1}^{\infty}(1-e^{\frac{2\pi i t}{M}}q^{\frac{m}{N\lambda}}y^{\ell})^{\hat{c}_{g,h}(0,m,\ell,t)}\Big) . \tag{3.54}$$

The numerator in (3.53) is the analogue of the “Hodge anomaly” in [50].

Infinite products of the form (3.51) were studied in [37], where it was shown that they converge for $\left(\begin{smallmatrix}\tau & z\\ z & \sigma\end{smallmatrix}\right)$ in a suitable domain of the Siegel upper half-space of 2×2 complex symmetric matrices with positive definite imaginary part

$$\mathbb{H}_2 := \Big\{Z\in \mathrm{Mat}_2(\mathbb{C}) \mid Z=Z^t,\ \mathrm{Im}Z>0\Big\} . \tag{3.55}$$

Furthermore, these products can be analytically continued to meromorphic functions on the whole $\mathbb{H}_2$, with zeroes and (possibly) poles along the rational quadratic divisor [58] (see section 4.2). As will be discussed in section 5.2, the $\Phi_{g,h}$ are Siegel modular forms under certain discrete subgroups of $Sp(4,\mathbb{R})$.

The forms $\Phi_{g,h}$ depend on the choice of the cocycle representative α. If α and α' differ by a 3-coboundary $\partial\beta$, then the corresponding forms $\Phi_{g,h}$ and $\Phi'_{g,h}$ are related by a shift in σ and an overall phase

$$\Phi'_{g,h}(\sigma,\tau,z) = e^{-2\pi i\nu_{g,h}}\Phi_{g,h}(\sigma+\nu_{g,h},\tau,z)\ , \tag{3.56}$$

where $\nu_{g,h}$ satisfies $e^{2\pi i\nu_{g,h}} = \frac{\beta(g,h)}{\beta(h,g)}$.

3.3.3 Multiplicative versus additive lift

The construction of a Siegel modular form $\Phi(\sigma,\tau,z)$ from a (weak) Jacobi form $\phi(\tau,z)$ via an infinite product representation, as exemplified by (3.51), is generally referred to as a *multiplicative (automorphic) lift*. The Jacobi form ϕ is said to be the *multiplicative seed* of the lift (see, for instance, [58, 50, 51, 59]), and we write

$$\Phi = \mathrm{Mult}[\phi]. \tag{3.57}$$

It should be stressed that the definition of $\Phi_{g,h}$ in this paper is based on the (α-twisted) equivariant Hecke operators $\mathcal{T}_L^\alpha$ rather than the ordinary T_L, as in the standard multiplicative lifts. As a consequence, the infinite product expression (3.51) involves the Fourier coefficients of many distinct Jacobi forms $\phi_{g^d,g^{-b}h^a}$, rather than a unique ϕ as, for example, in [58].

In some cases one can also obtain the Siegel modular form using a different procedure, known as the *additive lift*[‡‡]. In this case the Siegel modular form is constructed as a certain generating function (without exponentiation) of Hecke operators acting on a different Jacobi form ψ, which is then called the *additive seed*. For $\psi(\tau,z)$ a Jacobi form of weight k we write $\tilde{\psi}(\sigma,\tau,z) = p\psi(\tau,z)$ and define the additive lift as

$$\Phi = \mathrm{Add}[\psi] := \sum_{m\geq 1} m^{2-k}(T_m^-\tilde{\psi})(\sigma,\tau,z), \tag{3.58}$$

where T_m^- is a certain Hecke operator; see, e.g. [60] for the precise definition and properties of the right hand side.

As an example, consider the case of the Igusa cusp form Φ_{10} (corresponding to $\Phi_{e,e}$). The multiplicative seed is the K3 elliptic genus $\phi_{e,e} = \phi_{0,1}$ which is the unique weak Jacobi form of weight 0 and index 1. The multiplicative lift yields an infinite product formula for Φ_{10} [61]:

$$\Phi_{10} = \mathrm{Mult}[\phi_{0,1}] = pqy \prod_{(d,m,\ell)>0} \left(1-q^m y^\ell p^d\right)^{c(d,m,\ell)} \tag{3.59}$$

[‡‡] This generalizes the *Saito-Kurokawa-Maass lift* defined in [57].

where $c(d,m,\ell)$ are the Fourier coefficients of $\phi_{0,1}$. This is indeed obtained from (3.51) by restricting to $(g,h)=(e,e)$. As explained in [51] one can also obtain this Siegel modular form via an additive lift from the seed $\phi_{10,1}$ (the unique weak Jacobi form of weight 10 and index 1):

$$\Phi_{10} = \mathrm{Add}[\phi_{10,1}]. \tag{3.60}$$

We now observe that the additive seed $\phi_{10,1}$ can be expressed as $\vartheta(\tau,z)^2\eta(\tau)^{18}$ which is precisely the symmetrization factor $\psi_{g,h}$ in (3.54) when restricting to $(g,h)=(e,e)$. This is in fact a general feature that holds whenever a Siegel modular form can be realized both as an additive lift as well as a multiplicative lift. It is a consequence of the fact that the additive seed ψ is the first Fourier-Jacobi coefficient in the expansion of Φ and it is precisely this coefficient which appears as the prefactor in the symmetrization of the infinite product. For purposes we thus expect that for all (g,h) for which an additive lift exists we should have the following equalities

$$\Phi_{g,h} = \mathrm{Mult}[\phi_{g,h}] = \mathrm{Add}[\psi_{g,h}]. \tag{3.61}$$

We stress that one can of course define an additive lift of a Jacobi form via (3.58). However it is not guaranteed that when applying this procedure to $\psi_{g,h}$ one will always reproduce the functions $\Phi_{g,h}$. Moreover, not all automorphic infinite products have associated additive lifts. In general we expect that whenever $\Phi_{g,h}$ can be represented through an additive lift, the seed will be $\psi_{g,h}$. In section 5.3 we indeed verify this for a number of examples. For $(g,h)=(e,h)$ modified versions of these additive lifts have also been considered in [62].

4 Wall-crossing and Mason's generalized moonshine for M_{24}

Already in 1990, Mason proposed an M_{24} version of Norton's generalized moonshine conjecture [32, 33]. In this section, we will establish the connection between Mason's generalized moonshine for M_{24} and the recent Mathieu moonshine involving the $K3$ elliptic genus. This involves taking the multiplicative lift $\phi_{g,h}\to\Phi_{g,h}$, defined in section 3, after which the limit $z\to 0$ reproduces the generalized eta-products $\eta_{g,h}$ constructed by Mason [33]. Thus, the exponential lift $\Phi_{g,h}$ links Mason's generalized M_{24}-moonshine to the M_{24}/K3-moonshine considered here. This connection was first suggested in [2] for the special case of the twining genera $\phi_{1,h}=\phi_h$. A pictorial overview of this relation is given in figure 1.2. We begin by reviewing Mason's construction of the generalized eta-products $\eta_{g,h}$ and explain their interpretation in terms of partition functions of twisted chiral bosons. Physically, the Siegel modular forms $\Phi_{g,h}$ have interpretations as generating functions of twisted dyons in (CHL) orbifolds of $\mathcal{N}=4$ string theory, and the eta-products emerge as a result of a wall-crossing formula, generalizing the one written down in [2] for $\Phi_{e,h}$.

4.1 Mason's generalized M_{24}-moonshine

For each commuting pair $(g,h) \in M_{24}$, Mason associates a modular function on the upper half-plane satisfying the requirements posed by Norton in [19]. The starting point of Mason was the 24-dimensional permutation representation of M_{24} in which each element can be associated to a *cycle shape*, which describes the element as a product of cycles of permutations. For instance, in this representation the identity element is represented by the cycle shape 1^{24}, corresponding to the product of 24 identity permutations, while the elements in class $2A$ are represented by $1^8 2^8$, corresponding to the product of 8 identity permutations followed by 8 consecutive order 2 permutations.

Mason's functions were all given in terms of so called *eta-products*, namely products of Dedekind eta-functions. Suppose an order M element $h \in M_{24}$ has cycle shape

$$\prod_{\ell|M} \ell^{i(\ell)} \ , \tag{4.1}$$

on the 24-dimensional representation of M_{24}, for some integers $i(\ell)$. Then one can associate an eta-product to h by

$$\eta_h(\tau) := \prod_{\ell|M} \eta(\ell\tau)^{i(\ell)}. \tag{4.2}$$

Mason found a generalization of such eta-products associated to "generalized cycle shapes" labelled by commuting pairs (g,h) in M_{24}. We shall denote these generalized moonshine functions by $\eta_{g,h}(\tau)$.[§§] Let $g,h \in M_{24}$ be a pair of commuting elements of order $o(g)=N$ and $o(h)=M$ and consider their action in the standard 24-dimensional representation V of M_{24}. Let $v_1,\dots,v_{24}$ be a basis of simultaneous eigenvectors for g and h, relative to the eigenvalues $(e^{\frac{2\pi i r_i}{N}}, e^{\frac{2\pi i t_i}{M}})$, $i=1,\dots,24$, with $1 \le r_i \le N$ and $1 \le t_i \le M$. Then, the eta-product $\eta_{g,h}$ is defined as

$$\eta_{g,h}(\tau) = q^{\frac{1}{N\lambda}} \prod_{i=1}^{24} \prod_{n=0}^{\infty} (1 - e^{\frac{2\pi i t_i}{M}} q^{\frac{r_i}{N}+n}) \ , \tag{4.3}$$

where λ is the length of the shortest cycle of g. In particular, the eta-products η_h of equation (4.2) correspond to $(g,h)=(e,h)$. The products $\eta_{g,h}$ (or rather their inverse) can be interpreted as h-twining partition functions for 24 g-twisted chiral free bosons.

[§§] In [33] these functions were denoted by $f(g,h;\tau)$.

As proven in [32], the eta products $\eta_{g,h}$ satisfy the modular transformations

$$\upsilon_{g,h}\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(c\tau+d)^{-w}\eta_{g^d h^{-c},g^{-b}h^a}\left(\frac{a\tau+b}{c\tau+d}\right)=\eta_{g,h}(\tau)\ , \tag{4.4}$$

for some $\upsilon_{g,h}\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in U(1)$. The full list of Mason's generalized eta-products and the corresponding weights w can be found in Table 3. Furthermore, by definition, $\eta_{g,h}$ are invariant under conjugation in M_{24}

$$\eta_{kgk^{-1},khk^{-1}}(\tau)=\eta_{g,h}(\tau)\ . \tag{4.5}$$

4.2 Connecting the two moonshines via wall-crossing

The connection between the $K3/M_{24}$-moonshine for twining genera $\phi_h(\tau,z)$ and the M_{24}-moonshine associated with the functions $\eta_h(\tau)$ was pointed out in [2]. We now want to use our results above to establish this link also between the two different *generalized* moonshines for M_{24}.

As already mentioned, infinite products of the form (3.51) have been considered by Borcherds [58]. Although in general $\Phi_{g,h}$ are not expected to be Borcherds products in a strict sense [29], the same argument as in [58] shows that the infinite products (3.51) converge in some region of the Siegel upper half-space $\mathbb{H}_2$ and they can be analytically continued to meromorphic functions on the whole of $\mathbb{H}_2$. Furthermore, in the domain of absolute convergence of the infinite products (3.51), the zeroes and poles of $\Phi_{g,h}$ can only be located at rational quadratic divisors of the form:

$$H_{d,m,\ell,v}=\left\{\left(\begin{smallmatrix} \tau & z \\ z & \sigma \end{smallmatrix}\right) \mid \tfrac{m}{N\lambda}\tau+\ell z+d\sigma+\frac{v}{M}=0\right\}\ , \tag{4.6}$$

where $d,m,\ell,v\in\mathbb{Z}$ are coprime integers ($\gcd(d,m,\ell,v)=1$) that satisfy a positive discriminant condition

$$\ell^2-\frac{4md}{N\lambda}>0\ . \tag{4.7}$$

This property is an immediate consequence of the infinite product representation (3.51) of $\Phi_{g,h}$; the divisors of Borcherds products admit a similar description [37, 58, 63]. Using the fact that every point in the Siegel upper half-space can be brought into a domain of convergence using an $Sp(4;\mathbb{Z})$-transformation, and that the family of functions $\Phi_{g,h}$ is preserved under this action (see section 5), it follows that all divisor components can be brought to the form (4.6), for suitable $N\lambda, M$, by some $Sp(4;\mathbb{Z})$-transformation.[¶¶] This formula can be viewed as a generalisation of the rational quadratic divisor found in [25] in the context of twisted dyon counting in CHL-models.

[¶¶] Note that $Sp(4,\mathbb{Z})$ acts non-linearly on the variables σ,τ,z; therefore, even though the divisor (4.6) is defined by a linear equation in σ,τ,z, this is in general not true for its modular images.

The multiplicity of the zero or pole of $\Phi_{g,h}$ at the divisor $H_{d,m,\ell,v}$ is given by $\hat{c}_{g,h}(d,m,\ell,v)$. In particular, $\Phi_{g,h}$ is holomorphic if and only if $\hat{c}_{g,h}(d,m,\ell,v)$ is non-negative at every rational quadratic divisor. As will be shown in section 5, $\Phi_{g,h}$ is an automorphic form for a discrete subgroup $\Gamma^{(2)}_{g,h} \subset Sp(4,\mathbb{R})$. Therefore, one only needs to consider the distinct orbits of rational quadratic divisors under the action of $\Gamma^{(2)}_{g,h}$ to determine the full divisor of $\Phi_{g,h}$. The precise calculation of the divisor of each $\Phi_{g,h}$ requires knowledge of the modular group $\Gamma^{(2)}_{g,h}$ for each commuting pair (g,h) and will be left to future work.

A special role is played by (the modular orbit of) the divisor $H_{0,0,-1,0}$, corresponding to the locus $z=0$ in $\mathbb{H}_2$. Since $c_{e,g}(0,-1)=2$ for all $g\in M_{24}$, we have

$$\hat{c}_{g,h}(0,0,-1,0) = \frac{1}{MN\lambda}\sum_{k=0}^{M-1}\sum_{b=0}^{\lambda N-1} c_{e,g^{-b}h^k}(0,-1) = 2\ . \tag{4.8}$$

It follows that every $\Phi_{g,h}$ has a double zero at this divisor. Using the relation

$$\sum_{\ell\in\mathbb{Z}} c_{g,h}(r,\ell) = 0\ , \qquad \text{for } r>0\ , \tag{4.9}$$

together with the relations

$$\mathrm{Tr}_{\mathcal{H}_e(0,\pm1)}(g^{-b}h^k(-1)^{F+\bar{F}}) = 2\ , \tag{4.10}$$

$$\mathrm{Tr}_{\mathcal{H}_e(0,\pm\ell)}(g^{-b}h^k(-1)^{F+\bar{F}}) = 0 \qquad \text{for } \ell>1\ , \tag{4.11}$$

$$\sum_{\ell\in\mathbb{Z}}\mathrm{Tr}_{\mathcal{H}_e(0,\ell)}(g^{-b}h^k(-1)^{F+\bar{F}}) = \mathrm{Tr}_{\mathbf{24}}(g^{-b}h^k)\ , \tag{4.12}$$

we obtain

$$\begin{aligned}\lim_{z\to0}\frac{\Phi_{g,h}(\sigma,\tau,z)}{(2\pi iz)^2} &= \lim_{z\to0}\frac{y(1-y^{-1})^2}{(2\pi iz)^2}\Big(q^{\frac{1}{N\lambda}}\prod_{t=1}^{M}\prod_{m=1}^{\infty}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}})^{\sum_{\ell\in\mathbb{Z}}\hat{c}_{g,h}(0,m,\ell,t)}\Big)\\ &\qquad\times\Big(p\prod_{t=1}^{M}\prod_{d=1}^{\infty}(1-e^{\frac{2\pi it}{M}}p^d)^{\sum_{\ell\in\mathbb{Z}}\hat{c}_{g,h}(d,0,\ell,t)}\Big)\ .\end{aligned} \tag{4.13}$$

Next we use the relations (4.10)-(4.12) to find an expression for $\psi_{g,h}$ in (3.54) in terms of $\eta_{g,h}$ and known modular objects. Using the aforementioned equations we obtain

$$\begin{aligned}\psi_{g,h}(\tau,z) &= q^{\frac{1}{N\lambda}}y\prod_{t=0}^{M-1}\Big((1-e^{\frac{2\pi it}{M}}y^{-1})^{\hat{c}_{g,h}(0,0,-1,t)}\prod_{m=1}^{\infty}\prod_{\ell\in\mathbb{Z}}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}}y^{\ell})^{\hat{c}_{g,h}(0,m,\ell,t)}\Big)\\ &= q^{\frac{1}{N\lambda}}y(1-y^{-1})^2\prod_{n=1}^{\infty}\frac{(1-q^ny)^2(1-q^ny^{-1})^2}{(1-q^n)^4}\prod_{t=0}^{M-1}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}})^{\sum_{\ell\in\mathbb{Z}}\hat{c}_{g,h}(0,m,\ell,t)}.\end{aligned} \tag{4.14}$$

To proceed we note the identity

$$\begin{aligned}
&q^{\frac{1}{N\lambda}}\prod_{t=0}^{M-1}\prod_{m=1}^{\infty}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}})^{\sum_{k=0}^{M-1}\sum_{b=0}^{N\lambda-1}\frac{e^{-\frac{2\pi itk}{M}}}{M}\frac{e^{\frac{2\pi ibm}{N\lambda}}}{N\lambda}\mathrm{Tr}_{\mathbf{24}}(g^{-b}h^{k})}\\
&=q^{\frac{1}{N\lambda}}\prod_{t=0}^{M-1}\prod_{m=1}^{\infty}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}})^{\sum_{i=1}^{24}\frac{1}{M}\sum_{k=0}^{M-1}\frac{1}{N\lambda}\sum_{b=0}^{N\lambda-1}e^{-\frac{2\pi i(t-t_i)k}{M}}e^{\frac{2\pi ib(m-\lambda r_i)}{N\lambda}}}\\
&=q^{\frac{1}{N\lambda}}\prod_{i=1}^{24}\prod_{n=0}^{\infty}(1-e^{\frac{2\pi it_i}{M}}q^{n+\frac{r_i}{N}})=\eta_{g,h}(\tau)\ ,
\end{aligned}\tag{4.15}$$

where $1\le r_i\le N$ and $1\le t_i\le M$ are such that $(e^{\frac{2\pi ir_i}{N}},e^{\frac{2\pi it_i}{M}})$, $i=1,\dots,24$, are the (g,h)-eigenvalues of a basis of simultaneous eigenvectors for g and h in the 24-dimensional representation of M_{24}. Using this relation in (4.14) we obtain

$$\psi_{g,h}(\tau,z)=-\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^6}\eta_{g,h}(\tau)=\phi_{-2,1}(\tau,z)\eta_{g,h}(\tau)\ ,\tag{4.16}$$

where $\phi_{-2,1}$ is the standard weak Jacobi forms of weight -2 and index 1 for $SL(2,\mathbb{Z})$ (see appendix E).

To complete the analysis of the limit in (4.13) we note that

$$\lim_{z\to0}\frac{y(1-y^{-1})^2}{(2\pi iz)^2}\Big(q^{\frac{1}{N\lambda}}\prod_{t=1}^{M}\prod_{m=1}^{\infty}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N\lambda}})^{\sum_{\ell\in\mathbb{Z}}\hat{c}_{g,h}(m,0,\ell,t)}\Big)=\lim_{z\to0}\frac{\psi_{g,h}(\tau,z)}{(2\pi iz)^2}=\eta_{g,h}(\tau)\ .\tag{4.17}$$

For the pair $(g,h)=(e,e)$, the function $\Phi_{e,e}(\sigma,\tau,z)$ is invariant under the exchange $\sigma\leftrightarrow\tau$. From a physical viewpoint, this is a consequence of S-duality of type II superstring compactified on K3×T^2. In section 5.2, we will prove that $\Phi_{g,h}$ satisfy analogous transformations

$$\Phi_{g,h}(\sigma,\tau,z)=\Phi_{g,h'}(\frac{\tau}{N\lambda},N\lambda\sigma,z)\ ,\tag{4.18}$$

where $h'\in C_{M_{24}}(g)$ is not necessarily in the same conjugacy class as h. Using this identity, we conclude that

$$\lim_{z\to0}\frac{\Phi_{g,h}(\sigma,\tau,z)}{(2\pi iz)^2}=\eta_{g,h}(\tau)\eta_{g,h'}(N\lambda\sigma)\ .\tag{4.19}$$

We conjecture that this equation has a physical interpretation as a wall-crossing formula whenever $1/\Phi_{g,h}$ is the generating function of BPS-states in string theory. Indeed, for some pairs (g,h) we know that $\Phi_{g,h}^{-1}$ corresponds to the generating function for the h-twisted degeneracies of 1/4 BPS states in a CHL model. More precisely, these degeneracies are the Fourier coefficients of the automorphic form $\Phi_{g,h}^{-1}$ and the region where the Fourier expansion is performed depends on the moduli. These multiplicities jump as one crosses the pole at

$z = 0$; the physical interpretation is that some 1/4 BPS dyon corresponding to a bound state of 1/2 BPS configurations becomes unstable in a certain region of the moduli space and thus disappears from the spectrum. The mismatch $\eta_{g,h}(\tau)^{-1}\eta_{g,h'}(N\lambda\sigma)^{-1}$ between the Fourier coefficients at the two sides of the pole represents the degeneracy of a bound state of two 1/2 BPS states.

5 Automorphic Properties

In this final section we analyze the modular transformation properties of $\Phi_{g,h}(\sigma,\tau,z)$ with respect to discrete subgroups of $Sp(4,\mathbb{R})$. We show that they are Siegel modular forms and in some cases we are able to identify them with previously known objects. Our proof of modularity is done in two steps. The crucial first step is to determine the transformation properties of $\Phi_{g,h}(\sigma,\tau,z)$ with respect to the interchange $\sigma \leftrightarrow \tau$. We refer to this as "S-duality" since in the cases when $\Phi_{g,h}$ has an interpretation as a partition function in an $\mathcal{N} = 4$ string theory this corresponds precisely to the S-duality that exchanges electric and magnetic charges (see, e.g., [52, 64, 65, 24, 25]). It turns out that this transformation involves the subtle concept of "relabeling" of the group elements (g, h) introduced in [17]. By a careful analysis of this relabeling phenomenon we establish the S-duality symmetry in section 5.1. To complete the analysis of the automorphic properties of $\Phi_{g,h}(\sigma,\tau,z)$ we must also verify the transformation properties with respect to the remaining generators of the relevant modular subgroups of $Sp(4,\mathbb{R})$. This is rather straightforward since it essentially follows from the modularity of the seed functions $\phi_{g,h}(\tau,z)$ and $\psi_{g,h}(\tau,z)$. We present this analysis in section 5.2, and in section 5.3 we also investigate some examples in detail.

5.1 Relabeling and S-duality

In this section, we will show that the functions $\Phi_{g,h}(\sigma,\tau,z)$ defined by the infinite product (3.51) satisfy some 'S-duality' identities that exchange σ and τ. This property, together with the modular properties of the twisted twining genera $\phi_{g,h}$ proved in [17], will be sufficient to prove that all $\Phi_{g,h}$ are automorphic functions under some subgroup of $Sp(4,\mathbb{R})$ (see section 5.2).

5.1.1 Orbifolds and relabeling

In the case where g belongs to some M_{23} subgroup of M_{24}, the main step in the derivation of S-duality for $\Phi_{g,h}$ is the relabeling phenomenon [17]. This can be understood by considering the example of a holomorphic conformal field theory by a cyclic group. In this subsection, we review some of the standard properties of these orbifolds. See, for example, [53, 54, 39] for more details.

Let $\mathcal{C}$ be a holomorphic conformal field theory with automorphism group G. Given an element $g \in G$ of order N, let $\mathcal{C}'$ be the orbifold*** of $\mathcal{C}$ by $\langle g\rangle$. The spectrum $\mathcal{H}'$ of $\mathcal{C}'$ is constructed by considering the direct sum of the g^r-twisted representations $\mathcal{H}_{g^r}$ of $\mathcal{C}$, for all $r = 0, \dots, N-1$ and then restricting to the g-invariant sector, i.e.

$$\mathcal{H}' = \oplus_{r=0}^{N-1}\left(\mathcal{H}_{g^r}^{\langle g\rangle}\right) , \tag{5.1}$$

where $\mathcal{H}_{g^r}^{\langle g\rangle}$ denotes the g-invariant part of the g^r-twisted sector. The orbifold theory is a well-defined CFT, provided that one can define a consistent (in particularly, local) OPE between the g-invariant twisted fields. A necessary and sufficient consistency condition is that the cohomology class $[\alpha] \in H^3(G, U(1))$ is trivial when restricted to $H^3(\langle g\rangle, U(1))$ (level-matching). In what follows, we will assume that this condition is satisfied.

Each twisted sector carries a representation of some central extension $C_G^\alpha(g)$ of the centralizer $C_G(g)$ of g in G, that is compatible with the structure of twisted $\mathcal{C}$-module. As explained in section 3.2.3, there is an ambiguity in the definition of this representation. If the orbifold is consistent, the action of $C_G^\alpha(g)$ can be always chosen to be compatible also with the OPE of g-invariant twisted fields, so that it defines a group of automorphisms of the orbifold theory. In particular, for any $r, s \in \mathbb{Z}$, the representation of $C_G^\alpha(g)$ corresponding to the g^{r+s}-twisted sector must be contained in the tensor product of the representations corresponding to the g^r- and g^s-twisted sectors. A central extension satisfying this constraint can be chosen of the form

$$1 \to \langle Q\rangle \cong \mathbb{Z}_N \to C_G^\alpha(g) \to C_G(g) \to 1 , \tag{5.2}$$

where the central element Q (quantum symmetry) acts by $e^{\frac{2\pi i r}{N}}$ on the g^r-twisted sector.

Since the spectrum $\mathcal{H}'$ of $\mathcal{C}'$ consists of g-invariant states, the group acting faithfully on $\mathcal{H}'$ is $C_G^\alpha(g)/\langle g\rangle$. We conclude that $C_G^\alpha(g)/\langle g\rangle$ must be a subgroup of the group G' of automorphisms of the orbifold CFT $\mathcal{C}'$. For each $h \in C_G^\alpha(g)$, one can define the h-twining genera $Z'_{e,h}$ as in (2.3), such that

$$Z'_{e,h}(\tau) = \frac{1}{N}\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} Z_{g^r,g^sh}(\tau) . \tag{5.3}$$

***Here and in the following, we use the word ‘orbifold’ in the physicists’ sense, i.e. as a full-fledged two dimensional conformal field theory with modular invariant partition function.

Note that $Z'_{e,h}(\tau)$ only depends on the image of h in the projection $C^\alpha_G(g) \to C^\alpha_G(g)/\langle g\rangle \subseteq G'$.

The group G' of automorphisms of $\mathcal{C}'$ contains, in particular, an element g' corresponding to the quantum symmetry $g' \equiv Q$. It is well-known that by taking the orbifold of $\mathcal{C}'$ by the cyclic group $\langle g'\rangle$ one re-obtains the original CFT $\mathcal{C}$. More precisely, the g'^n-twisted sector $\mathcal{H}'_{g'^n}$ can be identified with the direct sum

$$\mathcal{H}'_{g'^n} = \oplus_{r=0}^{N-1} \mathcal{H}_{g^r,n} \ , \tag{5.4}$$

where $\mathcal{H}_{g^r,n}$ is the g-eigenspace with eigenvalue $e^{\frac{2\pi i n}{N}}$ in the g^r-twisted sector $\mathcal{H}_{g^r}$ of the CFT $\mathcal{C}$. Under the identification (5.4), the g'-invariant sector of the direct sum $\oplus_{n=0}^{N-1}\mathcal{H}'_{g'^n}$ corresponds indeed to the spectrum $\mathcal{H}$ of the original theory $\mathcal{C}$ and the quantum symmetry Q' corresponds to g. The identification (5.4) can be refined as

$$\mathcal{H}'_{g'^n,r} \cong \mathcal{H}_{g^r,n} \ , \tag{5.5}$$

where $\mathcal{H}'_{g'^n,r}$ is the g'-eigenspace with eigenvalue $e^{\frac{2\pi i m}{N}}$ in the twisted sector $\mathcal{H}'_{g'^n}$. Eq.(5.5) is an isomorphism between (untwisted) modules over the common subalgebra $\mathcal{C}^{\langle g\rangle} = \mathcal{C}'^{\langle g'\rangle}$ of $\mathcal{C}$ and $\mathcal{C}'$.

Each eigenspace $\mathcal{H}_{g^r,n}$ is a representation of $C^\alpha_G(g)$. Analogously, each twisted sector $\mathcal{H}'_{g'^n}$ (and, in fact, each eigenspace $\mathcal{H}'_{g'^n,r}$) carries an action of a suitable central extension

$$1 \to \langle Q'\rangle \to C^{\alpha'}_{G'}(g') \to C_{G'}(g') \to 1 \ , \tag{5.6}$$

of $C_{G'}(g')$ for some 3-cocycle α' representing a class in $H^3(G', U(1))$. In fact, there is an isomorphism $\varphi : C^\alpha_G(g) \xrightarrow{\cong} C^{\alpha'}_{G'}(g')$ such that $\varphi(Q) = g'$, $\varphi(g) = Q'$ and the identification (5.5) is equivariant with respect to the action of the corresponding groups. As a consequence, eq.(5.3) can be generalized to obtain the formulae for the twisted twining partition functions in the theory $\mathcal{C}'$

$$Z'_{g'^n,\varphi(h)}(\tau) = \frac{1}{N}\sum_{r,s\in\mathbb{Z}/N\mathbb{Z}} e^{-\frac{2\pi i n s}{N}} Z_{g^r,g^s h}(\tau) \ , \qquad h \in C^\alpha_G(g) \ . \tag{5.7}$$

A special case occurs when the two CFTs $\mathcal{C}'$ and $\mathcal{C}$ are isomorphic, so that $G' \cong G$. If, in addition, g' and g are in the same conjugacy class of G, the isomorphism φ can be chosen to be an outer automorphism of the group $C^\alpha_G(g)$ that exchanges g and Q, i.e. $\varphi(g) = Q$ and $\varphi(Q) = g$ and such that

$$\sum_{b=0}^{N-1} \frac{e^{\frac{2\pi i r b}{N}}}{N} Z_{g^m,g^{-b}\varphi(h)}(\tau) = \sum_{s=0}^{N-1} \frac{e^{\frac{2\pi i m s}{N}}}{N} Z_{g^r,g^{-s}h}(\tau) \ . \tag{5.8}$$

Since φ is outer, it might happen that $\varphi(h)$ and h belong to different conjugacy classes of $C^\alpha_G(g)$. If this is the case, we say that the class of h has been relabeled. Roughly speaking, a similar phenomenon occurs for the Mathieu moonshine. In particular, we will see that S-duality relates a function $\Phi_{g,h}$ to a function $\Phi_{g,h'}$, where h' belongs to the 'relabeled' class.

5.1.2 Relabeling for Mathieu Moonshine

The construction described in the previous section in the case of a holomorphic conformal field theory can be repeated with minor modifications for the abstract modules $\mathcal{H}_g$ underlying the generalized Mathieu Moonshine. It should be stressed, however, that since no consistent CFT with automorphism group M_{24} and spectrum $\mathcal{H}$ exists, these properties do not follow directly from the general theory of holomorphic CFTs and need to be proved independently.

Consider an element $g \in G = M_{24}$ of order N and let $\mathcal{H}_g$ be the g-twisted sector in the generalized Mathieu Moonshine conjecture. Let $\mathcal{H}'$ be the g-invariant subspace of the direct sum of all twisted sectors

$$\mathcal{H}' = \oplus_{r=0}^{N-1} \left(\mathcal{H}_{g^r}^{\langle \tilde{\rho}_{g,r}(g) \rangle} \right) , \tag{5.9}$$

where $\mathcal{H}_{g^r}^{\langle \tilde{\rho}_{g,r}(g) \rangle}$ denotes the $\tilde{\rho}_{g,r}(g)$-invariant part of the g^r-twisted sector. The level-matching condition, i.e. the requirement that the restriction of the cohomology class $[\alpha]$ is trivial in $H^3(\langle g \rangle, U(1))$, is satisfied if and only if g is contained in some M_{23} subgroup of M_{24}, so we will consider only this case. If $\mathcal{H}$ can be interpreted as the spectrum of R-R right-moving ground states in a non-linear sigma model on K3 with a symmetry g, then $\mathcal{H}'$ corresponds to the spectrum of right-moving ground states in the g-orbifold theory. The orbifold is a consistent $\mathcal{N} = (4,4)$ SCFT with central charge 6, which turns out to be again a non-linear sigma model on K3, since its elliptic genus is

$$\begin{aligned} \mathrm{Tr}_{\mathcal{H}'} \left(q^{L_0 - \frac{c}{24}} \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} y^{J_0^3} (-1)^{F+\bar{F}} \right) &= \sum_{r,s=0}^{N-1} \frac{1}{N} \mathrm{Tr}_{\mathcal{H}_{g^r}} \left(\tilde{\rho}_{g,r}(g)^s q^{L_0 - \frac{c}{24}} \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} y^{J_0^3} (-1)^{F+\bar{F}} \right) \\ &= \frac{1}{N} \sum_{r,s=1} \phi_{g^r,g^s}(\tau, z) = \phi_{e,e}(\tau, z) \ . \end{aligned} \tag{5.10}$$

Here, the twisted twining genera ϕ_{g^r,g^s} are relative to a cocycle α satisfying the conditions (D.17) and (D.21), in particular with trivial restriction to $\langle g \rangle$. As shown in appendix D.2, in this case the associated central extension $C^\alpha_{M_{24}}(g)$ is finite

$$1 \to \langle Q \rangle \cong \mathbb{Z}_N \to C^\alpha_{M_{24}}(g) \to C_{M_{24}}(g) \to 1 \ . \tag{5.11}$$

In the theory of orbifold CFTs, the conditions (D.17) and (D.21) ensure that the action of $C^\alpha_{M_{24}}(g)$ is compatible with the OPE of the twisted fields. The identity (5.10) can be verified case by case for all $g \in M_{24}$ satisfying the level-matching condition, and holds independently of the existence of a physical interpretation in terms of orbifolds of K3 sigma models.

By (5.10), the space $\mathcal{H}'$ defined by (5.9) is isomorphic to $\mathcal{H}$ as a module over the $\mathcal{N} = 4$ superconformal algebra. As a consequence, one can define a representation of $G' = M_{24}$ over $\mathcal{H}'$ satisfying the properties of Mathieu moonshine. Furthermore, $\mathcal{H}'$ also carries a representation ρ' of $C^{\alpha}_{M_{24}}(g)/\langle g_\alpha \rangle$ given by the restriction of $\oplus_{r=0}^{N-1} \tilde{\rho}_{g,r}$ to $\mathcal{H}'$. It can be proved that, as an abstract group, $C^{\alpha}_{M_{24}}(g)/\langle g_\alpha \rangle$ is isomorphic to a subgroup of M_{24} [17]. More precisely, Q is identified with an element $g' \in G' \cong M_{24}$ in the same conjugacy class as g and the image of $C^{\alpha}_{M_{24}}(g)/\langle g \rangle$ is the centralizer $C_{M_{24}}(g')$ of g'. Therefore, the 'orbifold' construction for the Mathieu moonshine modules is perfectly analogous to the case, described in the subsection 5.1.1, where a holomorphic CFT $\mathcal{C}$ and its orbifold $\mathcal{C}'$ are isomorphic.

In analogy with the subsection 5.1.1, we can introduce some g'^n-twisted sectors $\mathcal{H}'_{g'^n}$ satisfying (5.4) and (5.5), now considered as isomorphisms of $\mathcal{N} = 4$ modules. Each twisted sector $\mathcal{H}'_{g'^n}$ carries a representation of a central extension $C^{\alpha'}_{M_{24}}(g')$ defined as in eq.(5.6). It is natural to conjecture that there is an isomorphism $\varphi : C^{\alpha}_{M_{24}}(g) \xrightarrow{\cong} C^{\alpha'}_{M_{24}}(g')$ such that the identifications (5.5) are equivariant with respect to the action of the corresponding groups. Since $g \in G$ and $g' \in G'$ belong to the same conjugacy class of $G \cong G' \cong M_{24}$, they can be identified through a suitable choice of the isomorphism $G \cong G'$. With this identification, φ defines an outer automorphism of $C^{\alpha}_{M_{24}}(g)$ exchanging g and Q. Furthermore, we expect the analogue of (5.8) to hold. More precisely, we can state the following:

Conjecture 5.1. *For all $g \in M_{23} \subset M_{24}$ and all choices of cocycle α satisfying (D.17) and (D.21) and the corresponding central extension $C^{\alpha}_{M_{24}}(g)$ as in (5.11), there is an automorphism $\varphi : C^{\alpha}_{M_{24}}(g) \to C^{\alpha}_{M_{24}}(g)$ such that $\varphi(g) = Q$, $\varphi(Q) = g$, and for all $k \in C^{\alpha}_{M_{24}}(g)$, $m, r \in \mathbb{Z}$,*

$$\sum_{b=0}^{N-1} \frac{e^{\frac{2\pi i r b}{N}}}{N} \phi_{{g_\alpha}^m, {g_\alpha}^{-b} \varphi(k)}(\tau, z) = \sum_{s=0}^{N-1} \frac{e^{\frac{2\pi i m s}{N}}}{N} \phi_{g_\alpha^r, g_\alpha^{-s} k}(\tau, z) \ . \tag{5.12}$$

The existence of automorphisms that exchange g and Q is easy to check for the elements g of order higher than 4 (thus, excluding the classes 2A and 4B of M_{24}). Indeed, in all these cases the central extension of $C_{M_{24}}(g)$ has the form $C^{\alpha}_{M_{24}}(g) \cong \langle Q \rangle \times (\langle Qg \rangle.G)$. This group clearly admits an automorphism that fixes the factor $\langle Qg \rangle.G$ and exchanges Q and g. For the classes $2A$ and $4B$, the existence of such automorphisms can be verified with the aid of the software GAP (see also the ancillary files in the arXiv version of [17]).

In order to prove the conjecture, one needs to show that, for each g, one of these automorphisms φ satisfies (5.12). Heuristically, by standard CFT arguments, a conjecture of the form 5.1 and the corresponding identities (5.12) are expected to hold for any group $H \subseteq C_{M_{24}}(g)$

admitting an interpretation as a group of symmetries of a K3 sigma model. According to the analysis in [8], such an interpretation exists whenever $H \subset M_{24}$ has at least four orbits in the 24-dimensional permutation representation; this condition is satisfied by most of the non-cyclic abelian groups $\langle g, h\rangle \subset M_{24}$ (see Table 2 in appendix F).

A more rigorous and general proof can be obtained through a direct computation. To this aim, it is useful to rephrase eq.(5.12) in terms of identities among the twisted twining genera $\phi_{g,h}$ relative to the centralizer $C_{M_{24}}(g)$ rather than the genera ϕ_{g_α,h_α} relative to its central extension $C^\alpha_{M_{24}}(g)$. Recall that any element $k \in C^\alpha_{M_{24}}(g)$ can be written as $k = Q^x h_\alpha$, for some $x \in \mathbb{Z}/N\mathbb{Z}$, $h \in C_{M_{24}}(g)$. Therefore, one needs to show that, for each $h \in C_{M_{24}}(g)$, the image $\varphi(h_\alpha) = Q^{x'} h'_\alpha$ of the lift $h_\alpha \in C^\alpha_{M_{24}}(g)$ satisfies

$$e^{\frac{2\pi i m x'}{N}} \sum_{b=0}^{N-1} \frac{e^{\frac{2\pi i r b}{N}}}{N} \epsilon_{g,h'} \left(\begin{smallmatrix} 1 & b \\ 0 & m \end{smallmatrix}\right) \phi_{g^m, g^{-b}h'}(\tau, z) = \sum_{s=0}^{N-1} \frac{e^{\frac{2\pi i m s}{N}}}{N} \epsilon_{g,h} \left(\begin{smallmatrix} 1 & s \\ 0 & r \end{smallmatrix}\right) \phi_{g^r, g^{-s}h}(\tau, z) \ , \tag{5.13}$$

for all $m, r \in \mathbb{Z}$. Given the explicit knowledge of the twisted twining genera $\phi_{g,h}$, the proof of the conjecture then amounts to verifying a long series of identities between Jacobi forms.

We have verified that, for each conjugacy class of g, there is essentially only one choice for the conjugacy class h' and $x' \in \mathbb{Z}/N\mathbb{Z}$ for which all these identities can be possibly satisfied. In particular, if g is not in one of the classes $2A$ or $4B$, then the automorphism φ necessarily satisfies

$$h' = h \ , \qquad x' = 0 \ , \tag{5.14}$$

for all $h \in C_{M_{24}}(g) \setminus \langle g\rangle$. For g in one of the classes 2A or 4B, the conjugacy class of h' in $C_{M_{24}}(g)$ for each $C_{M_{24}}(g)$-conjugacy class of h is reported in Tables 4 and 5. Notice that, in general, h and h' might belong to different conjugacy classes of M_{24}; this phenomenon has been dubbed 'relabeling' in [17]. When h and h' are in different conjugacy classes of $C_{M_{24}}(g)$, we can choose the cocycle α in such a way that $x' = 0$. When h and h' are conjugated in $C_{M_{24}}(g)$, then x' can be set to 0 if and only if $\varphi(h_\alpha)$ is conjugated with h_α in the central extension $C^\alpha_{M_{24}}(g)$. This is always the case, except for g in M_{24}-class $4B$ and h in one of the classes $2\mathrm{B}_3$ or $4\mathrm{B}_1$ of $C_{M_{24}}(g)$. In both these cases, the pair g, h generates group 20 in the list of appendix F and $\varphi(h_\alpha)$ is conjugated with $Q^2 h_\alpha$, i.e. $x' = 2$.

We have proved a subset of the identities (5.13), required for the analysis in section 5.1.3. This leads to the following theorem:

Theorem 5.2. *Conjecture 5.1 holds for all g in the classes $1A$, $2A$, $3A$, $4B$ and $8A$ of M_{24}.*

The identities (5.13) can be easily verified by considering the modular properties of both sides and comparing a sufficient number of Fourier coefficients. This is a tedious but totally straightforward calculation, so we omit the details. We leave a systematic analysis of these identities and the complete proof of the conjecture to future work.

Notice that whenever a generator g of a non-cyclic abelian subgroup $\langle g, h\rangle \subset M_{24}$ is also an element of $M_{23} \subset M_{24}$, then it is necessarily in one of the classes 2A, 3A, 4B or 8A and theorem 5.2 applies (see appendix F for a list of all such groups and their generators). As will be explained in the subsection 5.1.3, the identities (5.13) are sufficient to establish the 'S-duality invariance' of the corresponding automorphic forms $\Phi_{g,h}$. The conjecture 5.1 is trivially true in the degenerate case where g is the identity, so the automorphic forms $\Phi_{e,h}$ associated with cyclic groups satisfy an analogous property [28, 29]. For the groups $\langle g, h\rangle$ that have no generator in M_{23}, the modular properties of $\Phi_{g,h}$ will be determined in a different way in section 5.1.3.

5.1.3 S-duality invariance

In this section, we will show that the functions $\Phi_{g,h}(\sigma, \tau, z)$ defined in terms of the infinite product (3.51), satisfy identities of the form

$$\Phi_{g,h}(\sigma, \tau, z) \sim \Phi_{g,h'}(\frac{\tau}{N\lambda} + x, N\lambda\sigma, z)\ , \tag{5.15}$$

for some suitable $h' \in C_{M_{24}}(g)$ and real x that depends on the cocycle α. Here, $\sim$ denotes equality up to a phase, that depends on the cocycle α. More precisely, for each abelian subgroup $\langle g, h\rangle \subset M_{24}$, we will prove identities of the form (5.15) for at least one pair of generators (g, h). Furthermore, in all such cases the cocycle α can be chosen in such a way that $x = 0$. For the remaining pairs $(g, h) \in \mathcal{P}$, the identities (5.15) can be derived through a complete analysis of the automorphic properties of $\Phi_{g,h}$, as discussed in section 5.2.

When the function $\Phi_{g,h}$ admits a physical interpretation as the generating functions for twisted multiplicities of 1/4 BPS states in a CHL model, the identity (5.15) corresponds to S-duality exchanging electric and magnetic charges in the low energy effective action.

Case 1: *g in one of the classes* 1A, 2A, 3A, 4B *or* 8A

Let g be an element of $M_{23} \subset M_{24}$ of order N in one of the classes considered in theorem 5.2. The restriction of $[\alpha]$ to $H^3(\langle g\rangle, U(1))$ is trivial, so that $\lambda = 1$ and we can choose a representative α satisfying (D.17) and (D.21). Consider an element $h \in C_{M_{24}}(g)$ and let M be the order of the lift $h_\alpha \in C^\alpha_{M_{24}}(g)$. Then eq.(5.13) implies that for any $d, m, k \in \mathbb{Z}_{\geq 0}$, there are $h' \in C_{M_{24}}(g)$ and $x' \in \mathbb{Z}/N\mathbb{Z}$ such that

$$\sum_{s=0}^{N-1} \frac{e^{\frac{2\pi i d s}{N}}}{N} \epsilon_{g,h}\left(\begin{smallmatrix} k & s \\ 0 & m \end{smallmatrix}\right) \phi_{g^m, g^{-s}h'^k}(\tau, z) = \sum_{b=0}^{N-1} \frac{e^{\frac{2\pi i m(b-kx')}{N}}}{N} \epsilon_{g,h}\left(\begin{smallmatrix} k & b \\ 0 & d \end{smallmatrix}\right) \phi_{g^d, g^{-b}h^k}(\tau, z)\ . \tag{5.16}$$

Thus,

$$\begin{aligned}
\hat{c}_{g,h}(d,m,\ell,t) &= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=0}^{N-1} \frac{e^{\frac{2\pi i b m}{N}}}{N} \epsilon_{g,h}\left(\begin{smallmatrix} k & b \\ 0 & d \end{smallmatrix}\right) c_{g^d,g^{-b}h^k}\left(\frac{md}{N},\ell\right) &(5.17)\\
&= \sum_{k=0}^{M-1} e^{\frac{2\pi i k m x'}{N}} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{s=0}^{N-1} \frac{e^{\frac{2\pi i d s}{N}}}{N} \epsilon_{g,h'}\left(\begin{smallmatrix} k & s \\ 0 & m \end{smallmatrix}\right) c_{g^m,g^{-s}h'^k}\left(\frac{md}{N},\ell\right) &(5.18)\\
&= \hat{c}_{g,h'}\left(m,d,\ell,t+mx'\frac{M}{N}\right) \ . &(5.19)
\end{aligned}$$

If h and h' are in different conjugacy classes of $C_{M_{24}}(g)$, we can choose the representative cocycle α in the class $[\alpha]$ in such a way that $x'=0$, so that

$$\begin{aligned}
\Phi_{g,h}(\sigma,\tau,z) =& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t=0}^{M-1} \left(1-e^{\frac{2\pi i t}{M}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h}(d,m,\ell,t)} &(5.20)\\
=& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t=0}^{M-1} \left(1-e^{\frac{2\pi i t}{M}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h'}(m,d,\ell,t)} &(5.21)\\
=& \Phi_{g,h'}\left(\frac{\tau}{N},N\sigma,z\right) \ . &(5.22)
\end{aligned}$$

When h and h' are in the same conjugacy class of $C_{M_{24}}(g)$, then x' cannot be eliminated in general. This corresponds to the case when $\varphi(h_\alpha)$ is conjugated with $Q^{x'}h_\alpha$ in the central extension $C^{\alpha}_{M_{24}}(g)$. In particular, since $\varphi(h_\alpha)^M=\varphi(h_\alpha^M)=e$, this implies that $Q^{x'}$ has order M, i.e. there is an integer x such that

$$x'M = xN \ . \tag{5.23}$$

Thus,

$$\begin{aligned}
\Phi_{g,h}(\sigma,\tau,z) =& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t=0}^{M-1} \left(1-e^{\frac{2\pi i t}{M}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h}(d,m,\ell,t)} &(5.24)\\
=& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t=0}^{M-1} \left(1-e^{\frac{2\pi i t}{M}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h}(m,d,\ell,t+mx)} &(5.25)\\
=& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t'=0}^{M-1} \left(1-e^{\frac{2\pi i t'}{M}} e^{-\frac{2\pi i x m}{M}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h}(m,d,\ell,t')} &(5.26)\\
=& pq^{\frac{1}{N}}y \prod_{(d,m,\ell)>0} \prod_{t'=0}^{M-1} \left(1-e^{\frac{2\pi i t'}{M}} e^{-\frac{2\pi i x' m}{N}} q^{\frac{m}{N}} y^{\ell} p^{d}\right)^{\hat{c}_{g,h}(m,d,\ell,t')} &(5.27)\\
=& e^{\frac{2\pi i x'}{N}} \Phi_{g,h}\left(\frac{\tau-x'}{N},N\sigma,z\right) \ . &(5.28)
\end{aligned}$$

In fact, $x' \neq 0$ only when g is in class 4B and h in one of the classes $2B_3$ or $4B_1$ of $C_{M_{24}}(g)$. In both cases, the pair g, h generates group 20 in the list of appendix F and $x' = 2$. For such pairs (g, h), however, since $T_L^\alpha \phi_{g,h} = 0$ for L odd, the function $\Phi_{g,h}$ satisfies the identity

$$\Phi_{g,h}(\sigma, \tau, z) = -\Phi_{g,h}(\sigma + \frac{1}{2}, \tau, z) \ . \tag{5.29}$$

Using this identity, we conclude

$$\Phi_{g,h}(\sigma, \tau, z) = \Phi_{g,h}(\frac{\tau}{N}, N\sigma, z) \ . \tag{5.30}$$

If the conjecture 5.1 holds, then an S-duality property holds for all $\Phi_{g,h}$ such that g is an element of $M_{23} \subset M_{24}$.

Case 2: *g in one of the classes $2B$, $3B$, $4C$, $6B$, $12B$*

The calculations in this case are rather technical and have therefore been relegated to appendix D.3. Below we summarize the results.

For a suitable choice of the representative 3-cocycle α, we have

$$\Phi_{g,h}(\sigma, \tau, z) = \Phi_{g,h^{-1}}(\frac{\tau}{N^2}, N^2\sigma, z) \ , \tag{5.31}$$

if h and h^{-1} are not conjugated within $C_{M_{24}}(g)$, and

$$\Phi_{g,h}(\sigma, \tau, z) = \Phi_{g,h}(\frac{\tau}{N^2}, N^2\sigma, z) \ , \tag{5.32}$$

otherwise.

5.2 Automorphic properties

Using the results of the previous sections, we can now prove that the functions $\Phi_{g,h}(\sigma, \tau, z)$, defined by the analytic continuation of the infinite product (3.51) to the Siegel upper half-space $\mathbb{H}_2$ are automorphic forms under certain subgroups of $Sp(4, \mathbb{R})$ acting by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot Z = (AZ + B)(CZ + D)^{-1} \ , \qquad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(4, \mathbb{R}) \tag{5.33}$$

where

$$Z = \begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} \ . \tag{5.34}$$

Recall that the function $\Phi_{g,h}$ can be written as

$$\Phi_{g,h}(\sigma, \tau, z) = p\psi_{g,h}(\tau, z) \exp[-\sum_{N=1}^{\infty} p^L (\mathcal{T}_L^\alpha \phi_{g,h})(\tau, z)] \ , \tag{5.35}$$

where the right-hand side converges on a suitable domain in $\mathbb{H}_2$.

Let us consider the action of various generators of $Sp(4,\mathbb{R})$ on $\Phi_{g,h}$.

- The Heisenberg subgroup $H(\mathbb{Z})$ of $Sp(4,\mathbb{Z})$ is defined as

  $$H(\mathbb{Z}) = \{[\zeta,\mu,\kappa] := \begin{pmatrix}1 & 0 & 0 & \mu\\ \zeta & 1 & \mu & \kappa\\ 0 & 0 & 1 & -\zeta\\ 0 & 0 & 0 & 1\end{pmatrix}, \zeta,\mu,\kappa \ \in \mathbb{Z}\}\ , \tag{5.36}$$

  and acts by

  $$[\zeta,\mu,\kappa]\cdot(\sigma,\tau,z) = (\sigma+\kappa+2\zeta z+\zeta^2\tau,\tau,z+\mu+\zeta\tau)\ . \tag{5.37}$$

  By (5.35), using the elliptic properties of the Jacobi forms $\psi_{g,h}$ and $\mathcal{T}_L^\alpha\phi_{g,h}$, we conclude easily that every $\Phi_{g,h}$ is invariant under $H(\mathbb{Z})$.

- For many commuting pairs $g,h\in M_{24}$, the genera $\mathcal{T}_L^\alpha\phi_{g,h}$ vanish unless L is an integer multiple of some $r\equiv r_{g,h}\in\mathbb{Z}$. Therefore, $\Phi_{g,h}$ is invariant up to a multiplier under $\sigma\mapsto\sigma+\frac{1}{r}$, i.e.

  $$\Phi_{g,h}([0,0,\frac{\kappa}{r}]\cdot(\sigma,\tau,z)) = e^{-\frac{2\pi i\kappa}{r}}\Phi_{g,h}(\sigma,\tau,z)\ , \tag{5.38}$$

  where $[0,0,\frac{\kappa}{r}]\cdot(\sigma,\tau,z) = (\sigma+\frac{\kappa}{r},\tau,z)$. In particular, $r_{g,h}=1$ when g,h generate a cyclic group and for groups 13, 23, 24 and 27, $r=3$ for groups 33 and 34, $r=4$ for groups 25 and 26 and $r=2$ in all the other cases. Notice that $r_{g,h}$ always divides $N\lambda$. More generally, the transformation $[0,0,\frac{\kappa}{N\lambda}]$ relates the functions $\Phi_{g,h}$ relative to two distinct choices of the cocycle α, that differ from one each other by a $N\lambda$-root of unity.

- The twisted twining genera $\mathcal{T}_L^\alpha\phi_{g,h}$ are Jacobi forms of weight 0 and index L under a group $\Gamma_{g,h}\subset SL(2,\mathbb{Z})$ (see the tables in appendix F), up to a multiplier $\chi_{g,h}^L$. As noticed above, for each group $\langle g,h\rangle$, there is an integer $r\equiv r_{g,h}$ such that $\mathcal{T}_L^\alpha\phi_{g,h}$ vanishes unless $r|L$. Therefore, only the power $\chi_{g,h}^r$ needs to be a well-defined character of $\Gamma_{g,h}$.

  On the other hand, the properties (4.4) and (4.5) imply that the eta products $\eta_{g,h}$ are modular forms of weight w and multiplier $\upsilon_{g,h}$ under the same group $\Gamma_{g,h}$ associated with $\phi_{g,h}$. Equivalently, each $\psi_{g,h}$ is a weak Jacobi form of weight $w-2$, index 1 and multiplier $\upsilon_{g,h}$ under $\Gamma_{g,h}$.

  For each $(\gamma,k)\in SL(2,\mathbb{Z})\times M_{24}$, let us choose $\mu_{g,h}(\gamma,k)\in\mathbb{R}/\mathbb{Z}$ such that

  $$e^{2\pi i r_{g,h}\mu_{g,h}(\gamma,k)} = \Big(\epsilon_{g,h}(\gamma,k)\Big)^{r_{g,h}}\ ,\qquad \gamma\in SL(2,\mathbb{Z})\ . \tag{5.39}$$

Then for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z})$ and $k \in M_{24}$, we have

$$\upsilon_{g,h}(\gamma)\Phi_{(\gamma,k)\cdot(g,h)}\Big(\xi(\gamma)\cdot Z\Big) = \det(c\tau+d)^{w-2}\,\Phi_{g,h}(\sigma+\mu_{g,h}(\gamma,k),\tau,z)\ , \tag{5.40}$$

where $Z = \left(\begin{smallmatrix} \tau & z \\ z & \sigma \end{smallmatrix}\right)$ and

$$\xi\begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ c & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\ , \tag{5.41}$$

so that

$$\xi(\gamma)\cdot\begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} = \begin{pmatrix} \frac{a\tau+b}{c\tau+d} & \frac{z}{c\sigma+d} \\ \frac{z}{c\sigma+d} & \sigma-\frac{z^2}{c\sigma+d} \end{pmatrix}\ . \tag{5.42}$$

In particular, by taking $(\gamma,k) \in M_{24}$ stabilizing (g,h), we obtain

$$\upsilon'_{g,h}(\gamma,k)\Phi_{g,h}\Big([0,0,-\mu_{g,h}(\gamma,k)]\cdot\big(\xi(\gamma)\cdot Z\big)\Big) = \det(c\tau+d)^{w-2}\,\Phi_{g,h}(\sigma,\tau,z)\ , \tag{5.43}$$

where $\gamma \in \Gamma_{g,h}$ and

$$\upsilon'_{g,h}(\gamma,k) = \upsilon_{g,h}(\gamma)e^{-2\pi i\mu_{g,h}(\gamma,k)}\ . \tag{5.44}$$

- Finally, for several pairs (g,h) of commuting elements of M_{24}, we have proved that

  $$\Phi_{g,h}(Z) = \Phi_{g,h'}(V_{N\lambda}\cdot Z)\ , \tag{5.45}$$

  for a suitable $h' \in C_{M_{24}}(g)$ and a suitable choice of the cocycle α, where

  $$V_t = \frac{1}{\sqrt{t}}\begin{pmatrix} 0 & t & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & t & 0 \end{pmatrix}\ , \tag{5.46}$$

  acts by

  $$V_t\cdot\begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} = \begin{pmatrix} t\sigma & z \\ z & \frac{\tau}{t} \end{pmatrix}\ . \tag{5.47}$$

  Notice that in each of the 55 conjugacy classes of abelian subgroups $\langle g,h\rangle \subset M_{24}$, an identity of the form (5.45) has been proved for at least one pair of generators.

As discussed in [51, 63], for any integers $N, t > 0$, the elements

$$V_t; \qquad \xi(\gamma)\ , \quad \gamma \in \Gamma_0(N); \qquad [\zeta,\mu,\kappa/t]\ , \quad \zeta,\mu,\kappa \in \mathbb{Z}\ , \tag{5.48}$$

generate the group $\Gamma_t^+(N) = \langle \Gamma_t(N), V_t\rangle \subset Sp(4,\mathbb{R})$ which is a normal double extension of the paramodular group

$$\Gamma_t(N) = \{ \begin{pmatrix} * & t* & * & * \\ * & * & * & t^{-1}* \\ N* & Nt* & * & * \\ Nt* & Nt* & t* & * \end{pmatrix} \in Sp(4,\mathbb{Q}),\ * \in \mathbb{Z}\}\ . \tag{5.49}$$

From the discussion above, it follows that every $\Phi_{g,h}$ is a modular function under some finite index subgroup $\Gamma^{(2)}_{g,h}$ of a paramodular group $\Gamma_t \equiv \Gamma_t(1)$, for some suitable t. The image of $\Phi_{g,h}$ under the action of a generic element of Γ_t is expected to be again a function $\Phi_{g',h'}$ for some (possibly different) commuting pair $g', h' \in M_{24}$, and defined with respect to a suitable choice of the cocycle α.

Under the action of $Sp(4,\mathbb{Z})$, the functions $\Phi_{g,h}(\sigma,\tau,z)$, for all commuting $g, h \in M_{24}$, are mapped into one another, up to shifts and rescalings of the arguments, i.e., schematically, $\Phi_{g,h}(\gamma\cdot(\sigma,\tau,z)) \sim \Phi_{g',h'}(\sigma/t + x, t\tau, z)$ for suitable $t \in \mathbb{Z}$, $x \in \mathbb{R}$ (we dropped the automorphy factors). In particular, $\Phi_{g,h}$ is a Siegel modular form for some congruence subgroup of $Sp(4,\mathbb{Z})$, as proved in the following theorem.

Theorem 5.3. *For each pair of commuting elements $g, h \in M_{24}$ and for a suitable choice of the cocycle α, the function $\Phi_{g,h}(Z)$ is a meromorphic Siegel modular function of weight $w-2$ and level H for a suitable H, i.e. it is a meromorphic function on the upper half space $\mathbb{H}_2$ such that*

$$\Phi_{g,h}\big((AZ+B)(CZ+D)^{-1}\big) = \det(CZ+D)^{w-2}\Phi_{g,h}(Z)\ , \tag{5.50}$$

for all $\left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right) \in \Gamma^{(2)}(H) := \ker(Sp(4,\mathbb{Z}) \to Sp(4,\mathbb{Z}/H\mathbb{Z}))$. Here, w is the weight of the corresponding Mason's eta function $\eta_{g,h}$.

Proof. It is sufficient to prove the statement for those 'special pairs' (g,h) for which we proved that $\Phi_{g,h}(\sigma,\tau,z) = \Phi_{g,h'}(\tau/t, t\sigma, z)$ for a suitable choice of representative cocycle α. Here, h' is some element in $C_{M_{24}}(g)$ (possibly different from h) and $t = N\lambda$, where, as usual, N denotes the order of g and λ the length of the shortest cycle of g as a permutation of 24 objects. Every other commuting pair of M_{24} elements can be obtained by an $SL(2,\mathbb{Z})$ transformation $(g,h)\cdot\gamma$ of one such 'special pair' (g,h), so that $\Phi_{(g,h)\cdot\gamma}$ is related to $\Phi_{g,h}$ by an $Sp(4,\mathbb{Z})$ transformation $\xi(\gamma)$. Thus, if $\Phi_{g,h}$ is a modular function under $\Gamma^{(2)}(H)$, then $\Phi_{(g,h)\cdot\gamma}$ is a modular function under $\xi(\gamma)\Gamma^{(2)}(H)\xi(\gamma)^{-1} = \Gamma^{(2)}(H)$.

There is some integer n such that the functions $\mathcal{T}_L^\alpha\phi_{g,h}$, $L \in \mathbb{Z}_{>0}$, and $\psi_{g,h}$ are weak Jacobi forms of weight 0 and index L (respectively, $w-2$ and 1) with trivial multiplier under the congruence subgroup $\Gamma(n) := \ker(SL(2,\mathbb{Z}) \to SL(2,\mathbb{Z}/n\mathbb{Z}))$. We can choose n so that also $\mathcal{T}_L^{\alpha}\phi_{g,h'}$ and $\psi_{g,h'}$ are weak Jacobi forms with trivial multiplier under the same group. Note that n is necessarily a multiple of $t = N\lambda$, since this is the smallest integer for which $\phi_{g,h}(\tau + N\lambda, z) = \phi_{g,h}(\tau, z)$. Thus, both $\Phi_{g,h}(Z)$ and $\Phi_{g,h'}(Z) = \Phi_{g,h}(V_t \cdot Z)$ transform as modular forms of weight $w-2$ under $\xi(\Gamma(n))$ and $[\zeta,\mu,0]$, for all $\zeta,\mu \in \mathbb{Z}$. Equivalently, $\Phi_{g,h}$ transforms as a Siegel modular form of weight $w-2$ under the group generated by

$$\xi(\Gamma(n)), \qquad V_t\xi(\Gamma(n))V_t, \qquad [\zeta,\mu,0], \qquad V_t[\zeta,\mu,0]V_t \qquad \zeta,\mu \in \mathbb{Z}\ , \tag{5.51}$$

which is a subgroup of $Sp(4,\mathbb{Z})$, since $t|n$. We will prove that this group contains $\Gamma^{(2)}(H)$, with $H = nt$. By Theorem 12.4 and Proposition 13.2 of [66], $\Gamma^{(2)}(nt)$ is generated by $\xi(\Gamma(nt))$ together with matrices of the form $\left(\begin{smallmatrix} \mathbb{I} & ntB \\ 0 & \mathbb{I}\end{smallmatrix}\right)$, $\left(\begin{smallmatrix} \mathbb{I} & 0 \\ ntC & \mathbb{I}\end{smallmatrix}\right)$, with B, C arbitrary 2×2 integral symmetric matrices. It is easy to check that every element of the form $\left(\begin{smallmatrix} \mathbb{I} & ntB \\ 0 & \mathbb{I}\end{smallmatrix}\right) \in \Gamma^{(2)}(nt)$ is contained in the group generated by $[\zeta,\mu,0]$ and $\xi\left(\begin{smallmatrix} 1 & bn \\ 0 & 1\end{smallmatrix}\right)$, with $\zeta,\mu,b \in \mathbb{Z}$, and every element of the form $\left(\begin{smallmatrix} \mathbb{I} & 0 \\ ntC & \mathbb{I}\end{smallmatrix}\right) \in \Gamma^{(2)}(nt)$ is contained in the group generated by $\xi\left(\begin{smallmatrix} 1 & 0 \\ cn & 1\end{smallmatrix}\right)$, $V_t\xi\left(\begin{smallmatrix} 1 & 0 \\ cn & 1\end{smallmatrix}\right)V_t$ and $V_t[\zeta,0,0]\xi\left(\begin{smallmatrix} 1 & 0 \\ n & 1\end{smallmatrix}\right)[-\zeta,0,0]V_t$, for all $\zeta, c \in \mathbb{Z}$. Clearly $\xi(\Gamma(nt))$ is a subgroup of $\xi(\Gamma(n))$, and this concludes the proof. $\square$

In section 5.3, we will discuss the precise automorphic properties of some of these functions $\Phi_{g,h}$. We leave the detailed description of the groups $\Gamma^{(2)}_{g,h}$ for all pairs g,h to a future work.

5.3 Examples

Having established that the $\Phi_{g,h}$ are indeed Siegel modular forms we now wish to analyze some specific examples in detail. We denote the pairs (g,h) by the M_{24}-conjugacy class of g and the $C_{M_{24}}(g)$-conjugacy class of h. See appendix F for more details.

$\Phi_{2A,2A_{2,3,5}}$ **(Groups 1, 2, 3)**

Groups 1, 2 and 3 are $\mathbb{Z}_2 \times \mathbb{Z}_2$ groups that contain three elements in class 2A of M_{24}. The corresponding $\Phi_{g,h}$ are identical

$$\Phi_{2A,2A_2}(\sigma,\tau,z) = \Phi_{2A,2A_3}(\sigma,\tau,z) = \Phi_{2A,2A_5}(\sigma,\tau,z)\ . \tag{5.52}$$

The 3-cocycle α can be chosen to have trivial restriction to these groups. The Hecke transforms of the twisted twining genera $\mathcal{T}_L\phi_{g,h}$ vanish unless L is a multiple of $r = 2$; in this case, they are Jacobi forms under $\Gamma_{g,h} = SL(2,\mathbb{Z})$ with trivial multiplier $\chi^2_{g,h} = 1$. The eta products for these elements are given by

$$\eta_{g,h}(\tau) = \eta(\tau)^{12}\ , \tag{5.53}$$

and are modular forms of weight $w = 6$ under $SL(2,\mathbb{Z})$ with multiplier (see eq. (5.40))

$$\upsilon_{g,h}(T) = -1 \qquad \upsilon_{g,h}(S) = -1 \ . \tag{5.54}$$

Furthermore, from the results of section 5.1.3 we deduce that, since $g \in M_{23}$ (with $N = 2$ and $\lambda = 1$), $\Phi_{g,h}$ satisfies

$$\Phi_{2A,2A_2}(\sigma,\tau,z) = \Phi_{2A,2A_3}(\frac{\tau}{2}, 2\sigma, z) \ , \qquad \Phi_{2A,2A_5}(\sigma,\tau,z) = \Phi_{2A,2A_5}(\frac{\tau}{2}, 2\sigma, z) \ . \tag{5.55}$$

Thus, $\Phi_{g,h}$ is a modular form of weight $w - 2 = 4$ with a multiplier $\upsilon_{g,h}$ given above under the subgroup of $Sp(4,\mathbb{R})$ generated by the Heisenberg group $H(\mathbb{Z})$, by $SL(2,\mathbb{Z})$ and under V_2. These elements of $Sp(4,\mathbb{R})$ generate the paramodular group $\Gamma_2^+(1)$. Furthermore, $\Phi_{g,h}$ has a double zero at the rational quadratic divisor $z \to 0$ and at all modular images of this divisor. This allows to identify $\Phi_{g,h}$ as

$$\Phi_{g,h}(Z) = \Delta_2(Z)^2 \ , \tag{5.56}$$

where Δ_2 is the modular form of weight 2 under $\Gamma_2^+(1)$ defined by Gritsenko and Nikulin in [51]. Notice that in this case the function $\psi_{g,h}$ in (4.16) is given by

$$\psi_{g,h}(\tau,z) = -\vartheta_1(\tau,z)^2\eta(\tau)^6, \tag{5.57}$$

which is the additive seed for $\Delta_2(Z)$ [51]. Hence we conclude that in this case

$$\Phi_{g,h} = \mathrm{Add}[\psi_{g,h}], \tag{5.58}$$

as claimed in section 3.3.3. The Siegel modular form $(\Delta_2)^2$ also appears in the context of umbral moonshine, where it corresponds to the Siegel modular form $\Phi^{(3)}(Z)$ (see section 2.6 of [30]). In this context $\Phi^{(3)}$ arises as the multiplicative lift of the umbral Jacobi form $Z^{(3)}(\tau,z)$ of weight 0 and index 2. It is therefore interesting to ask how this relates to our construction of $\Phi_{g,h}$ as the multiplicative lift (3.51). By comparing Fourier coefficients one can verify that we have the identity (see the table in appendix F for the result of the Hecke action)

$$\mathcal{T}_2\phi_{g,h}(\tau,z) = Z^{(3)}(\tau,z). \tag{5.59}$$

However, this does not necessarily imply that the two lifts are the same, since a priori our multiplicative lift in (3.51) differ from that in [30] since we are constructing it from the seed function $\phi_{g,h}$, which is a weak Jacobi form of weight 0 and index 1, using the equivariant Hecke operator $\mathcal{T}_L$ (the twist by the 3-cocycle α is trivial in this case). Even so, by virtue of (5.58), in the case at hand it turns out that the lifts do coincide and we thus have

$$\mathrm{Mult}[\phi_{g,h}] = \Phi^{(3)}. \tag{5.60}$$

$\Phi_{2A,2A_4}$ **(Group 7)**

The function $\Phi_{2A,2A_4}$ of group 7 satisfies

$$\Phi_{2A,2A_4}(V_2 \cdot Z) = \Phi_{2A,4B_1}(Z) \ , \tag{5.61}$$

where h' in class $4B_1$ of $C_{M_{24}}(g)$ is such that $h'^2 = g$. Therefore, the right hand side is a $Sp(4,\mathbb{Z})$ transformation of $\Phi_{e,4B}$, which is a Siegel modular form under $\Gamma_1(4) = \Gamma_0^{(2)}(4)$ [28, 29].

$\Phi_{2A,2B_{1,2}}$ **(Groups 8,9)**

The functions $\Phi_{g,h}$ for groups 8 and 9 are identical

$$\Phi_{2A,2B_1}(Z) = \Phi_{2A,2B_2}(Z) \ . \tag{5.62}$$

The twisted twining genera $\mathcal{T}_L^\alpha \phi_{g,h}$ vanish unless L is a multiple of $r = 2$; in this case, they are Jacobi forms under $\Gamma_{g,h} = \Gamma_0(2)$ with trivial multiplier. The eta products are given by

$$\eta_{g,h}(\tau) = \eta(\tau)^4 \eta(2\tau)^4 \ , \tag{5.63}$$

and are modular forms of weight $w = 4$ under $\Gamma_0(2)$ with multiplier

$$\upsilon_{g,h}\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = -1 \qquad \upsilon_{g,h}\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right) = -1 \ . \tag{5.64}$$

Therefore, $\Phi_{g,h}$ transforms as a modular form of weight $w - 2 = 2$ under $\xi(\gamma)$, $\gamma \in \Gamma_0(2)$, with multiplier $\upsilon_{g,h}(\gamma)$. Furthermore, for a suitable choice of the cocycle α, we have

$$\Phi_{2A,2B_{1,2}}(V_2 \cdot Z) = \Phi_{2A,2B_{1,2}}(Z) \ . \tag{5.65}$$

We conclude $\Phi_{g,h}$ is a modular form of weight 2 under the paramodular group $\Gamma_2(2)$. Assuming that it is holomorphic, then it can be identified as

$$\Phi_{g,h}(Z) = Q_1(Z)^2 \ , \tag{5.66}$$

where Q_1 is the modular form of weight 1 defined by Gritsenko and Clery in [63]. This conjecture is also supported by the fact that in this case $\psi_{g,h}$ is given by

$$\psi_{g,h}(\tau, z) = -\frac{\vartheta_1(\tau, z)^2}{\eta(\tau)^2} \eta(2\tau)^4, \tag{5.67}$$

which is the square of the additive seed for Q_1 (see eq. (16) in [63]).

$\Phi_{2A,4B_4}$ **(Group 12)**

The function $\Phi_{2A,4B_4}$ of group 12 satisfies

$$\Phi_{2A,4B_4}(V_2\cdot Z)=\Phi_{2A,8A_1}(Z)\ , \tag{5.68}$$

where h' in class $8A_1$ of $C_{M_{24}}(g)$ is such that $h'^4=g$. Therefore, the right hand side is a $Sp(4,\mathbb{Z})$ transformation of $\Phi_{e,8A}$, which is a Siegel modular form under $\Gamma_1(8)=\Gamma_0^{(2)}(8)$ [29].

$\Phi_{2A,4B_{2,3,5}}$ (Groups 17,18,19)

The functions $\Phi_{g,h}$ for groups 17, 18, 19 are identical

$$\Phi_{2A,4B_2}(Z)=\Phi_{2A,4B_3}(Z)=\Phi_{2A,4B_5}(Z)\ . \tag{5.69}$$

The Jacobi forms $\mathcal{T}_L^{\alpha}\phi_{g,h}$ for these groups have exactly the same modular properties as the ones for groups 8 and 9; the eta products are also the same. Therefore, $\Phi_{g,h}$ is a modular function of weight 2 under the paramodular group $\Gamma_2(2)$.

$\Phi_{4B,4B_{4,7}}$ (Groups 25,26)

The functions $\Phi_{g,h}$ for groups 25 and 26 are identical

$$\Phi_{4B,4B_4}(Z)=\Phi_{4B,4B_7}(Z)\ . \tag{5.70}$$

The 3-cocycle α can be chosen to be trivial when restricted to $\langle g,h\rangle$. The Hecke-transformed twisted twining genera $\mathcal{T}_L\phi_{g,h}$ vanish unless L is a multiple of $r=4$; in this case, they are Jacobi forms under $\Gamma_{g,h}=SL(2,\mathbb{Z})$ with trivial multiplier. The eta products are given by

$$\eta_{g,h}(\tau)=\eta(\tau)^6\ , \tag{5.71}$$

and are modular forms of weight $w=3$ under $SL(2,\mathbb{Z})$ with multiplier

$$\upsilon_{g,h}\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)=-i\qquad \upsilon_{g,h}\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)=-i\ . \tag{5.72}$$

Furthermore,

$$\Phi_{4B,4B_4}(V_4\cdot Z)=\Phi_{4B,4B_7}(Z)\ . \tag{5.73}$$

We conclude $\Phi_{g,h}$ is a modular function of weight $w-2=1$ under the paramodular group $\Gamma_4^+(1)$ and the multiplicity at the rational quadratic divisor $z\to 0$ is non-negative. It follows that

$$\Phi_{g,h}(Z)=\Delta_{1/2}(Z)^2\ , \tag{5.74}$$

where $\Delta_{1/2}$ is the modular form of weight 1/2 defined in [51]. Also in this case the additive lift matches since

$$\psi_{g,h}(\tau,z)=-\vartheta_1(\tau,z)^2 \tag{5.75}$$

is the square of the additive seed for $\Delta_{1/2}(Z)$, and hence $\mathrm{Add}[\psi_{g,h}] = (\Delta_{1/2}(Z))^2$.

The function $(\Delta_{1/2}(Z))^2$ also corresponds to the umbral Siegel modular form $\Phi^{(5)}(Z)$, which is the multiplicative lift of the umbral Jacobi form $Z^{(5)}(\tau, z)$ of weight 0 and index 4 [30]. The relation between $Z^{(5)}(\tau, z)$ and the twisted twining genus $\phi_{g,h}$ is now:

$$(\mathcal{T}_4\phi_{g,h})(\tau, z) = Z^{(4)}(\tau, z). \tag{5.76}$$

Again, it is not a priori clear that the equivariant multiplicative lift of $\phi_{g,h}$, defined by (3.51), will coincide with the ordinary multiplicative Borcherds lift of $Z^{(5)}(\tau, z)$ considered in [30]. However, in the case at hand they do:

$$\mathrm{Mult}[\phi_{g,h}] = \Phi^{(5)}. \tag{5.77}$$

$\Phi_{3A,3A_3}$ (Group 33)

Consider the function $\Phi_{3A,3A_3}$ of groups 33. The twisted twining genera $\mathcal{T}_L^{\alpha}\phi_{g,h}$ vanish unless L is a multiple of $r = 3$; in this case, they are Jacobi forms under $\Gamma_{g,h} = SL(2,\mathbb{Z})$ with trivial multiplier. The eta product is given by

$$\eta_{g,h}(\tau) = \eta(\tau)^8 \ , \tag{5.78}$$

and is a modular form of weight $w = 4$ under $SL(2,\mathbb{Z})$ with multiplier

$$\upsilon_{g,h}\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = e^{-\frac{2\pi i}{3}} \qquad \upsilon_{g,h}\left(\begin{smallmatrix} 0 & -1 \\ 0 & 1 \end{smallmatrix}\right) = 1 \ . \tag{5.79}$$

Furthermore,

$$\Phi_{3A,3A_3}(V_3 \cdot Z) = \Phi_{3A,3A_3}(Z) \ . \tag{5.80}$$

We conclude $\Phi_{g,h}$ is a modular function of weight $w - 2 = 2$ under the paramodular group $\Gamma_3^+(1)$ and the analysis of the multiplicities at its divisors shows that it is holomorphic with a double zero at $z = 0$. It follows that

$$\Phi_{3A,3A_3}(Z) = \Delta_1(Z)^2 \ , \tag{5.81}$$

where Δ_1 is the modular form of weight 1 defined in [51]. Also here we find that

$$\psi_{g,h}(\tau, z) = -\vartheta_1(\tau, z)^2\eta(\tau)^2 \tag{5.82}$$

is the square of the additive seed for $\Delta_1(Z)$. The function $\Delta_1(Z)^2$ coincides with the umbral Siegel modular form $\Phi^{(4)}(Z)$, which is the multiplicative lift of the umbral Jacobi form

$Z^{(4)}(\tau, z)$ of weight 0 and index 3. In this case we have that, for a suitable choice of cocycle α,

$$(\mathcal{T}_3^\alpha \phi_{g,h})(\tau, z) = Z^{(3)}(\tau, z), \tag{5.83}$$

and the equivariant and multiplicative lifts coincide:

$$\text{Mult}[\phi_{g,h}] = \Phi^{(4)}. \tag{5.84}$$

$\Phi_{3A,3B_1}$ (Group 34)

Consider the function $\Phi_{3A,3B_1}$ of group 34. The twisted twining genera $\mathcal{T}_L^\alpha \phi_{g,h}$ vanish unless L is a multiple of $r = 3$; in this case, they are Jacobi forms under $\Gamma_{g,h} = \Gamma_0(3)$ with trivial multiplier. The eta product is given by

$$\eta_{g,h}(\tau) = \eta(\tau)^2 \eta(3\tau)^2 \ , \tag{5.85}$$

and is a modular form of weight $w = 2$ under $SL(2, \mathbb{Z})$ with multiplier

$$\upsilon_{g,h}\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = e^{-\frac{2\pi i}{3}} \qquad \upsilon_{g,h}\left(\begin{smallmatrix} 1 & 0 \\ 3 & 1 \end{smallmatrix}\right) = e^{\frac{2\pi i}{3}} \ . \tag{5.86}$$

Furthermore,

$$\Phi_{3A,3B_1}(V_3 \cdot Z) = \Phi_{3A,3B_1}(Z) \ . \tag{5.87}$$

We conclude $\Phi_{g,h}$ is a modular function of weight $w - 2 = 0$ under the paramodular group $\Gamma_3^+(3)$. Since the weight vanishes, it must necessarily be meromorphic.

6 Conclusions

In this paper we have proposed a second quantized version of (generalized) Mathieu moonshine, involving a class of Siegel modular forms $\Phi_{g,h}$ for discrete subgroups $\Gamma_{g,h}^{(2)} \subset Sp(4, \mathbb{R})$, constructed from a multiplicative lift of the twisted twining genera $\phi_{g,h}$. For certain pairs of conjugacy classes of M_{24} we were able to identify $\Phi_{g,h}$ with known Siegel modular forms. It would be interesting to extend these results and perform a more detailed investigation of all the modular groups $\Gamma_{g,h}^{(2)}$ and determine whether the remaining $\Phi_{g,h}$ coincide with known objects, or perhaps constitute new examples of Siegel modular forms. In [29], Raum proved modularity for most of the cases $\Phi_{e,h}$ and found that not all of them are of a standard Borcherds product type, but in fact correspond to certain rescaled products of Borcherds modular forms. One would like to extend this analysis to determine whether the $\Phi_{g,h}$ for $g \neq e$ also contain such rescaled Borcherds products, or some generalization thereof.

As already mentioned in the introduction, an interesting by-product of our analysis is the fact that some of the Siegel modular forms $\Phi_{g,h}$ coincide with multiplicative lifts of the umbral Jacobi forms analyzed in [30]. This might be a simple consequence of the constraints from modularity, but it might also indicate some deeper relation between umbral moonshine and generalized Mathieu moonshine which would be interesting to uncover. Could it be that some of the other products $\Phi_{g,h}$ also coincide with lifts of the more general D_n or E_n Niemeier-umbral Jacobi forms in [67]?

It is natural to wonder whether our products $\Phi_{g,h}$ have interpretations in terms of denominator formulas of some generalized Kac-Moody algebras (GKMs). It is well-known that the inverse square root of the Igusa cusp form $\Phi_{10} = \Phi_{e,e}$ constitutes one side of the denominator formula for a rank 3 GKM-algebra of hyperbolic type [61] (the other side corresponds to the additive lift in (3.60)). In the terminology of Borcherds [37], the other functions $\Phi_{e,h}$ then correspond to twisted denominator formulas for the same algebra. On the other hand, by analogy with generalized Monstrous moonshine [42, 44], we would expect that the functions $\Phi_{g,e}$ give denominator formulas for a class of rank 3 GKM-algebras. For some of the elements g of small order these algebras have indeed been constructed in [68, 69, 70, 71] in the context of CHL-models (see also [61, 60, 51] for earlier mathematical results). In this context the prefactor $pq^{\frac{1}{N\lambda}}y$ in (3.51) should have an interpretation as the exponential of the Weyl vector ρ of the algebra. If true one would expect that for fixed g, the associated twisted denominator formulas $\Phi_{g,h}$ all have the same prefactor involving the Weyl vector of the original GKM-algebra determined by the class $[g]$. In other words, the prefactor should be independent of the twining element h, and this is indeed what we find. In fact, the Weyl vectors extracted from $pq^{\frac{1}{N\lambda}}y$ reduces to the ones in [68, 69] when the length of the shortest cycle λ equals one.

As stressed in the introduction, we think that our results could have immediate applications to the understanding of dyon counting in CHL-orbifolds. In particular, most of the functions $\Phi_{g,h}$ are expected to have interpretations as partition functions of twisted dyons, and we hope to investigate this relation in more detail in a future publication.

Acknowledgments

We are especially grateful to Matthias Gaberdiel for collaboration in the initial stages of this work, and for many helpful discussions. We are also grateful to Geoff Mason for several useful discussions on M_{24} moonshine for eta-products, and for sending us a printed copy of [32]; to Miranda Cheng for a discussion that drew our attention to the connection with umbral moonshine; and to Martin Raum for helpful discussions on infinite product representations of Siegel modular forms. In addition we thank Scott Carnahan, Thomas Creutzig, Terry Gannon, Suresh Govindarajan, Jeff Harvey, Stefan Hohenegger, Gerald Höhn, Jim Lepowsky, Greg Moore, Sameer Murthy, Anne Taormina and Katrin Wendland for inspiring discussions during the course of this work. We also thank the Simons Center and the organizers of the programme "Mock Modular Forms, Moonshine, and String Theory" for providing a stimulating research environment, and all the participants for many inspiring discussions. D.P. thanks the Albert Einstein Institute in Golm, and R.V. thanks the Department of Fundamental Physics at Chalmers for hospitality while this work was carried out.

A Some group cohomology

In this appendix, we summarize some general results about group cohomology. See [72] for more details.

For a finite group G, a 2-cochain $\beta : G \times G \to U(1)$ is closed (and hence defines a cocycle) provided it satisfies

$$\beta(g_1, g_2 g_3)\beta(g_2, g_3) = \beta(g_1 g_2, g_3)\beta(g_1, g_2) \tag{A.1}$$

for $g_1, g_2, g_3 \in G$. The second cohomology $H^2(G, U(1))$ then consists of the closed 2-cochains, modulo the ambiguity

$$\beta(g_1, g_2) \to \beta(g_1, g_2)\frac{\gamma(g_1)\gamma(g_2)}{\gamma(g_1 g_2)} , \tag{A.2}$$

where $\gamma : G \to U(1)$ is an arbitrary 1-cochain, i.e. an arbitrary function $\gamma : G \to U(1)$.

A 3-cochain α,

$$\alpha : G \times G \times G \to U(1) \tag{A.3}$$

is closed provided it satisfies

$$\alpha(g_1, g_2, g_3)\,\alpha(g_1, g_2 g_3, g_4)\,\alpha(g_2, g_3, g_4) = \alpha(g_1 g_2, g_3, g_4)\,\alpha(g_1, g_2, g_3 g_4) . \tag{A.4}$$

In the cohomology group $H^3(G, U(1))$ closed 3-cochains are then identified modulo

$$\alpha(g_1, g_2, g_3) \to \alpha(g_1, g_2, g_3)\,\frac{\beta(g_1 g_2, g_3)\beta(g_1, g_2)}{\beta(g_1, g_2 g_3)\beta(g_2, g_3)} . \tag{A.5}$$

Note that the multiplying factor is trivial if β is closed, i.e. if it satisfies the 2-cocycle condition (A.1). In particular, for each cocycle α and element $x \in G$, the 3-cocycle α_x, defined by

$$\alpha_x(g, h, k) := \alpha(x^{-1}gx,\ x^{-1}hx,\ x^{-1}kx) , \tag{A.6}$$

differs from α just by a 3-coboundary [47]

$$\alpha(g_1, g_2, g_3) = \alpha_x(g_1, g_2, g_3)\, \frac{\eta_x(g_1 g_2, g_3)\eta_x(g_1, g_2)}{\eta_x(g_1, g_2 g_3)\eta_x(g_2, g_3)}, \tag{A.7}$$

where

$$\eta_z(x, y) := \frac{\alpha(x, y, z)\alpha(z, z^{-1}xz, z^{-1}yz)}{\alpha(x, z, z^{-1}yz)} \ . \tag{A.8}$$

Given a 3-cocycle α, we can define, for any $h \in G$, a map $c_h : G \times G \to U(1)$ via

$$c_g(h_1, h_2) = \frac{\alpha(g, h_1, h_2)\alpha(h_1, h_2, (h_1 h_2)^{-1} g (h_1 h_2))}{\alpha(h_1, h_1^{-1} g h_1, h_2)} \ . \tag{A.9}$$

It is shown in [45] that c_g defines a 2-cocycle of the stabilizer subgroup $C_G(g) \subseteq G$ (i.e. the subgroup of all elements h_1, h_2 which commute with g). When $h_1, h_2 \in C_G(g)$, we have the simplified expression

$$c_g(h_1, h_2) = \frac{\alpha(g, h_1, h_2)\alpha(h_1, h_2, g)}{\alpha(h_1, g, h_2)} \ , \qquad h_1, h_2 \in C_G(g) \ . \tag{A.10}$$

It is straightforward to check this by inserting (A.9) into (A.1), and making repeated use of the 3-cocycle condition (A.4) (twice on each side of the equality sign) together with the fact that g commutes with h_1, h_2, h_3.

Under the 'gauge transformation' (A.5), c_g transforms as

$$c_g(h_1, h_2) \to \tilde{c}_g(h_1, h_2) := c_g(h_1, h_2)\frac{\gamma_g(h_1)\gamma_g(h_2)}{\gamma_g(h_1 h_2)} \ , \qquad h_1, h_2 \in C_G(g) \ , \tag{A.11}$$

where we defined the 1-cochain γ_g by

$$\gamma_g(h) \equiv \frac{\beta(g, h)}{\beta(h, g)} \ . \tag{A.12}$$

This is indeed of the form (A.2), and hence, for all $g \in G$, c_g defines a map

$$c_g \ : \ H^3(G, U(1)) \ \to H^2(C_G(g), U(1)) \ . \tag{A.13}$$

In fact, if c_g is the 2-cocycle associated to a projective representation ρ_g of $C_G(g)$, i.e.

$$\rho_g(h_1)\rho_g(h_2) = c_g(h_1, h_2)\rho_g(h_1 h_2) \ , \tag{A.14}$$

then $\tilde{c}_g$ is the 2-cocycle associated with the projectively equivalent representation

$$\tilde{\rho}_g(h) := \gamma_g(h)\rho_g(h) \ . \tag{A.15}$$

In the context of holomorphic CFTs, the transformation (A.5) corresponds to a redefinition (A.15) of the projective representations ρ_g of the centralizer $C_G(g)$ over the twisted sector $\mathcal{H}_g$, which induces the analogous transformation of the twisted twining partition functions

$$Z_{g,h} \to \tilde{Z}_{g,h} = \gamma_g(h) Z_{g,h} \ . \tag{A.16}$$

Indeed, the new partition functions $\tilde{Z}_{g,h}$ satisfy the expected modular properties with respect to the new cocycle $\tilde{\alpha}$.

In particular, it the case (A.7) of conjugation by $x \in G$, we have $\beta \equiv \eta_x^{-1}$, so that

$$\gamma_g(h) = \frac{\eta_x(h,g)}{\eta_x(g,h)} = \frac{c_g(x, x^{-1}hx)}{c_g(h,x)} \ , \qquad h \in C_g(G) \ , \tag{A.17}$$

where the latter equality follows from [46, 47]. From this identity, one recovers

$$Z_{(\mathbb{I},x^{-1})\cdot(g,h)}(\tau) \equiv Z_{x^{-1}gx, x^{-1}hx}(\tau) = \frac{c_g(x, x^{-1}hx)}{c_g(h,x)} Z_{g,h}(\tau) \ , \tag{A.18}$$

from which the formula (B.8) for $\epsilon_{g,h}(\mathbb{I}, x)$ follows.

One further useful identity is [47]

$$\frac{c_{x_1 x_2}(z_1, z_2)}{c_{x_1}(z_1,z_2) c_{x_2}(z_1,z_2)} = \frac{c_{z_1}(x_1,x_2) c_{z_2}(x_1,x_2)}{c_{z_1 z_2}(x_1,x_2)} \ , \tag{A.19}$$

that holds for *pairwise commuting* $x_1, x_2, z_1, z_2 \in G$.

For most applications, only the restriction of a 3-cocycle α to subgroups of the form $\langle g, h\rangle \cong \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}$ is needed. For such groups, we have $\mathcal{H}^3(\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}, U(1)) = \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2} \times \mathbb{Z}_{\gcd(N_1,N_2)}$ and a set of normalized representatives for the generators are [73]

$$\alpha_{v_i}(g^{a_1}h^{a_2}, g^{b_1}h^{b_2}, g^{c_1}h^{c_2}) := e^{\frac{2\pi i v_i a_i}{N_i^2}([b_i]_{N_i} + [c_i]_{N_i} - [b_i + c_i]_{N_i})} \tag{A.20}$$

where $v_i \in \mathbb{Z}/N_i\mathbb{Z}$, $i = 1, 2$, and

$$\alpha_{v_{12}}(g^{a_1}h^{a_2}, g^{b_1}h^{b_2}, g^{c_1}h^{c_2}) := e^{\frac{2\pi i v_{12} a_1}{N_1 N_2}([b_2]_{N_2} + [c_2]_{N_2} - [b_2 + c_2]_{N_2})} \tag{A.21}$$

where $v_{12} \in \mathbb{Z}/\gcd(N_1,N_2)\mathbb{Z}$ and $[\cdot]_x : \mathbb{Z} \to \{0, \ldots, x-1\}$ denotes the reduction modulo x. Notice that, with this choice for the generators, we have the simplified formula

$$c_x(y,z) = \alpha(x,y,z) \ , \qquad x, y, z \in \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2} \ . \tag{A.22}$$

B Modular properties of twisted twining partition functions

In this appendix we include some details on the modular properties of twisted twining partition functions. In particular, we analyze the combined action of $SL(2,\mathbb{Z})$ and G on the set of commuting pairs (g,h) in G. Here we must also take into account the presence of a non-trivial 3-cocycle α which leads to a certain twisted action. We introduce a convenient "twisted equivariant slash-operator" that simplifies many expressions since it combines (twisted) $SL(2,\mathbb{Z})$-equivariance with G-equivariance. Finally, we discuss a reformulated version of the cohomological obstructions found in [17].

B.1 $SL(2,\mathbb{Z})\times G$ - action on $\mathcal{P}_G$

It is clear from (2.8) that modular transformations act on $\tau\in\mathbb{H}$ as well as on the set of commuting pairs $(g,h)\in G\times G$. In addition G acts on itself by conjugation and thereby on the set of pairs g,h. In order to determine the subgroups $\Gamma_{g,h}\subset SL(2,\mathbb{Z})$ under which $Z_{g,h}$ (and $\phi_{g,h}$) are invariant for fixed $g,h\in G$, we must classify the orbits of the combined $SL(2,\mathbb{Z})\times G$-action on the commuting pair (g,h). We are mainly interested in the case $G=M_{24}$ but we only make this specification at the end of the subsection.

For any finite group G, let $\mathcal{P}_G\subset G\times G$ be the set of commuting pairs of elements:

$$\mathcal{P}_G=\{(g,h)\in G\times G\mid gh=hg\}\ . \tag{B.1}$$

The group $SL(2,\mathbb{Z})\times G$ has a left action as a permutation over this set by

$$\left(\begin{pmatrix}a&b\\c&d\end{pmatrix},k\right)\cdot(g,h):=(k\,g\,k^{-1},k\,h\,k^{-1})\begin{pmatrix}a&b\\c&d\end{pmatrix}^{-1}=(kg^dh^{-c}k^{-1},kg^{-b}h^ak^{-1})\ , \tag{B.2}$$

where $(g,h)\in\mathcal{P}_G$, $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in SL(2,\mathbb{Z})$ and $k\in G$. This action can be extended to an action of $GL(2,\mathbb{Z})\times G$ in the obvious way.

B.2 Twisted action on $\mathcal{P}_G$: generalized permutation

As explained in section 2.2, in order to include the possibility of non-trivial multipliers in the modular properties of the twisted twining genera one needs to consider the "twisting" of the action (B.2) by a 3-cocycle α, representing a cohomology class in $H^3(G,U(1))$. This α-twisted action is a generalized permutation

$$\mathbb{C}\mathcal{P}_G\times SL(2,\mathbb{Z})\times G\to\mathbb{C}\mathcal{P}_G\ , \tag{B.3}$$

on the complex vector space $\mathbb{C}\mathcal{P}_G$ freely generated by the elements of $\mathcal{P}_G$. We define this action by the formula

$$(\gamma, k)_\alpha : (g, h) \mapsto \epsilon_{g,h}(\gamma, k)(kgk^{-1}, khk^{-1})\gamma^{-1} \ . \tag{B.4}$$

Here, $(\gamma, k) \in SL(2,\mathbb{Z}) \times G$ and $\epsilon_{g,h}(\gamma, k) \in U(1)$ is a phase which depends on the choice of 3-cocycle α. More precisely, in terms of the 2-cocycle $c_g(h_1, h_2)$ in (2.7) the phases $\epsilon_{g,h}(\gamma, k)$ are defined as

$$\epsilon_{g,h}(\gamma_1\gamma_2, k_1k_2) := \epsilon_{(\gamma_2,k_2)\cdot(g,h)}(\gamma_1, k_1)\ \epsilon_{g,h}(\gamma_2, k_2) \ , \tag{B.5}$$

$$\epsilon_{g,h}(T, e) := \frac{1}{c_g(g, g^{-1}h)} \ , \tag{B.6}$$

$$\epsilon_{g,h}(S, e) := c_g(h^{-1}, h) \ , \tag{B.7}$$

$$\epsilon_{g,h}(\mathbb{I}, k) := \frac{c_g(h, k^{-1})}{c_g(k^{-1}, khk^{-1})} \ , \qquad k \in G \ , \tag{B.8}$$

where

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad\qquad T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \ , \tag{B.9}$$

are the generators of $SL(2,\mathbb{Z})$.

B.3 Modular properties of twisted twining partition functions

We now wish to analyze the modular properties of the twisted twining partition function $Z_{g,h}$ in more detail. To this end we define the following equivariant slash operator

$$f(g, h; \tau)|(\gamma, k) := f\big((\gamma, k) \cdot (g, h);\ \gamma \cdot \tau\big) \ , \qquad (\gamma, k) \in SL(2,\mathbb{Z}) \times G \ , \tag{B.10}$$

acting on modular functions $f : \mathcal{P}_G \times \mathbb{H}_+ \to \mathbb{C}$. We can think of the twisted twining partition functions $Z_{g,h}(\tau)$ as functions on $\mathcal{P}_G \times \mathbb{H}_+$ that are equivariant under $SL(2,\mathbb{Z}) \times G$, up to a multiplier. In the simplest case when the multiplier system is trivial, this property can be expressed as follows in terms of the slash operator:

$$Z_{g,h}(\tau)|(\gamma, k) = Z_{g,h}(\tau) \ . \tag{B.11}$$

In particular, for each fixed $(g, h) \in \mathcal{P}_G$, the function $Z_{g,h}(\tau)$ is a modular function under some subgroup $\Gamma_{g,h} \in PSL(2,\mathbb{Z})$, i.e.

$$Z_{g,h}(\gamma \cdot \tau) = Z_{g,h}(\tau) \ , \qquad \gamma \in \Gamma_{g,h} \subseteq PSL(2,\mathbb{Z}) \ . \tag{B.12}$$

The group $\Gamma_{g,h}$ is the image $\pi(\tilde{\Gamma}_{g,h})$ of the stabilizer

$$\tilde{\Gamma}_{g,h} := \{(\gamma, k) \in SL(2,\mathbb{Z}) \times G \mid (g,h) \cdot (\gamma, k) = (g,h)\} \ , \tag{B.13}$$

under the homomorphism

$$\pi : SL(2,\mathbb{Z}) \times G \to PSL(2,\mathbb{Z}) \ . \tag{B.14}$$

Similar properties hold for the twisted twining genera $\phi_{g,h}(\tau, z)$ of an $\mathcal{N} = (4,4)$ superconformal algebra: these are expected to be Jacobi forms of weight zero and index 1 with respect to the same groups $\Gamma_{g,h}$. For this reason, we need to extend the definition of the slash operator to an action on the space of functions $\psi : \mathcal{P}_G \times \mathbb{H}_+ \times \mathbb{C} \to \mathbb{C}$. Specifically, for Jacobi forms of weight 0 and index m with respect to some $\Gamma \subset SL(2,\mathbb{Z})$ we define

$$\psi(g,h;\tau,z)|\big(\gamma,k\big) := e^{-\frac{2\pi i m c z^2}{c\tau+d}} \psi\Big(k\, g^d h^{-c}\, k^{-1},\ k\, g^{-b} h^a\, k^{-1};\ \gamma\cdot\tau, \frac{z}{c\tau+d}\Big) \ , \tag{B.15}$$

where $(\gamma, k) = \big(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), k\big) \in SL(2,\mathbb{Z}) \times G$. Note that for theories with $\mathcal{N} = (4,4)$ superconformal symmetry, these genera are expected to be *even* functions of z, so that the central element $C = S^2 \in SL(2,\mathbb{Z})$ acts trivially and it makes sense to consider the action on such functions of the quotient $PSL(2,\mathbb{Z}) = SL(2,\mathbb{Z})/\mathbb{Z}_2$.

We are now ready to incorporate the α-twist in the modular properties of the twisted twining genera. To this end we define the α-twisted generalization of the equivariant slash operator:

$$f(g,h,\tau)|_\alpha(\gamma,k) := \epsilon_{g,h}(\gamma,k) f\big((\gamma,k)\cdot(g,h);\gamma\cdot\tau\big) \ , \qquad (\gamma,k) \in SL(2,\mathbb{Z}) \times G \ , \tag{B.16}$$

for some 3-cocycle α representing a class $[\alpha] \in H^3(G, U(1))$. Similarly, we define the α-twisted slash operators on Jacobi forms of weight 0 and index m by

$$\psi(g,h;\tau,z)|_\alpha\big(\gamma,k\big) := \epsilon_{g,h}(\gamma,k) e^{-\frac{2\pi i m c z^2}{c\tau+d}} \psi\Big(k\, g^d h^{-c}\, k^{-1},\ k\, g^{-b} h^a\, k^{-1};\ \gamma\cdot\tau, \frac{z}{c\tau+d}\Big) \ , \tag{B.17}$$

where $(\gamma, k) = \big(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), k\big) \in SL(2,\mathbb{Z}) \times G$.

B.4 Cohomological obstructions

When α represents a non-trivial class in $H^3(G, U(1))$, the partition functions $Z_{g,h}(\tau)$ are modular functions under $\Gamma_{g,h} \subseteq PSL(2,\mathbb{Z})$ only up to some multiplier $\chi_{g,h}$, which depends on α (see section 2.2). In fact, by (B.13) and (B.8), the restriction of $\epsilon_{g,h}$ to $\tilde{\Gamma}_{g,h}$ is a group homomorphism $\tilde{\Gamma}_{g,h} \to U(1)$. If

$$\epsilon_{g,h}(\gamma,k) = 1 \ , \tag{B.18}$$

for all (γ, k) in

$$\begin{aligned} \tilde{\Gamma}_{g,h} \cap \ker \pi = &\{(e,k) \in SL(2,\mathbb{Z}) \times G \mid (k^{-1}gk, k^{-1}hk) = (g,h)\} && \text{(B.19)} \\ &\cup \{(S^2,k) \in SL(2,\mathbb{Z}) \times G \mid (k^{-1}gk, k^{-1}hk) = (g^{-1},h^{-1})\} \ , && \text{(B.20)} \end{aligned}$$

then the restriction $\epsilon_{g,h} : \tilde{\Gamma}_{g,h} \to U(1)$ induces a well-defined homomorphism $\chi_{g,h} : \Gamma_{g,h} \to U(1)$ on the image $\Gamma_{g,h} = \pi(\tilde{\Gamma}_{g,h})$. Explicitly, for each $\gamma \in \Gamma_{g,h} \subseteq PSL(2,\mathbb{Z})$, one can choose a lift $(\gamma, k) \in \tilde{\Gamma}_{g,h} \subseteq SL(2,\mathbb{Z}) \times G$ and set

$$\chi_{g,h}(\gamma) := \epsilon_{g,h}(\gamma, k) \ , \qquad \text{for } \gamma = \pi(\gamma, k) \in \Gamma_{g,h} \ . \tag{B.21}$$

By (B.18), the definition is independent of the lift.

On the contrary, if (B.18) is not satisfied for some $(\gamma, k) \in \tilde{\Gamma}_{g,h} \cap \ker \pi$, then eq.(2.12) implies that $Z_{g,h}(\tau)$ must vanish identically. In this case, we say that the twisted twining partition function is *obstructed*. Following the discussion in [17, 18], we can distinguish between two kinds of obstructions:

(1) We say that, for a certain $(g,h) \in \mathcal{P}_G$, there is an obstruction of the *first kind* if there is some element of the form $(e,k) \in \tilde{\Gamma}_{g,h} \cap \ker \pi$ for which (B.18) is not satisfied.

(2) We say that there is an obstruction of the *second kind* if (B.18) is satisfied for all elements of the form $(e,k) \in \tilde{\Gamma}_{g,h} \cap \ker \pi$, but is violated by some element of the form $(S^2,k) \in \tilde{\Gamma}_{g,h} \cap \ker \pi$. By (B.8), this implies that (B.18) is false for *all* elements of the form $(S^2,k) \in \tilde{\Gamma}_{g,h} \cap \ker \pi$.

Although not phrased this way in [17] these obstructions are equivalent to the ones given there.

C Definition of twisted equivariant Hecke operators

In this appendix, we shall discuss the definition and properties of the α-twisted equivariant Hecke operators.

C.1 Twisted equivariant Hecke operator

Let us denote by $\mathrm{Mat}(\mathbb{Z})$ the ring of 2×2 integral matrices with positive determinant, graded by the determinant,

$$\mathrm{Mat}(\mathbb{Z}) = \bigcup_{L>0} \mathrm{Mat}_L(\mathbb{Z}) \ , \tag{C.1}$$

where $\mathrm{Mat}_L(\mathbb{Z})$ was defined in (3.4). For any $u \in \mathrm{Mat}_L(\mathbb{Z})$ denote by $u^\vee \in \mathrm{Mat}_L(\mathbb{Z})$ the dual

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\vee := L \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} , \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_L(\mathbb{Z}) \ . \tag{C.2}$$

Note that $\mathrm{Mat}_1(2,\mathbb{Z}) \equiv SL(2,\mathbb{Z})$ is the group of invertible elements of $\mathrm{Mat}(\mathbb{Z})$ and that each $\mathrm{Mat}_L(\mathbb{Z})$ is a bimodule over $SL(2,\mathbb{Z})$.

In order to define the α-twisted equivariant Hecke operators $\mathcal{T}_L^\alpha$, we need to extend the action (B.3) of $SL(2,\mathbb{Z}) \times G$ on $\mathbb{C}\mathcal{P}_G$ to an action

$$\mathbb{C}\mathcal{P}_G \times \mathrm{Mat}(\mathbb{Z}) \times G \to \mathbb{C}\mathcal{P}_G \ , \tag{C.3}$$

by

$$(u,k)_\alpha : (g,h) \mapsto \epsilon_{g,h}(u,k)(kgk^{-1}, khk^{-1})u^\vee \ , \tag{C.4}$$

for a suitable $\epsilon_{g,h}(u,k) \in U(1)$, that reduces to the one discussed in section B.2 when $u \in \mathrm{Mat}_1(\mathbb{Z}) = SL(2,\mathbb{Z})$. Furthermore, the interpretation of $\phi_{g,h}$ as a section of $\mathcal{L}^\alpha_{g,h}$ suggests that the following composition law should be imposed

$$(u_1u_2, k_1k_2)_\alpha \cdot (g,h) = (u_1,k_1)_\alpha \cdot \big((u_2,k_2)_{\alpha^L} \cdot (g,h)\big) \ , \tag{C.5}$$

where $(u_1,k_1) \in \mathrm{Mat}_L(\mathbb{Z}) \times G$ and $(u_2,k_2) \in \mathrm{Mat}(\mathbb{Z}) \times G$. In terms of the phases $\epsilon_{g,h}$, this condition reads

$$\epsilon_{g,h}((u_1,k_1)\cdots(u_n,k_n)) = \prod_{i=1}^n \epsilon_{(u_{i+1},k_{i+1})\cdots(u_n,k_n)\cdot(g,h)}(u_i,k_i)^{|u_1\cdots u_{i-1}|} \ , \tag{C.6}$$

where $|u_i| := \det u_i$.

Observe that $\mathrm{Mat}(\mathbb{Z})$, as a multiplicative semigroup, is generated by $SL(2,\mathbb{Z})$ together with the matrices of the form $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for p prime. Therefore, it is sufficient to specify the phases $\epsilon((\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}), e)$ for all primes p and any other phase $\epsilon_{g,h}(u,k)$, $(u,k) \in \mathrm{Mat}(\mathbb{Z}) \times G$, is then determined by (C.6). The α-twisted slash operator (2.15) can be trivially extended to an action with respect to $\gamma \in \mathrm{Mat}_L(\mathbb{Z})$, in which case it maps a Jacobi form ϕ of weight 0 and index m to another Jacobi form $\phi|_\alpha(\gamma,k)$ of weight 0 and index Lm. Notice that, by (C.6), the twisted slash operators satisfy the composition relation

$$|_\alpha(u_1u_2, k_1k_2) = |_\alpha(u_1,k_1)|_{\alpha^L}(u_2,k_2) \tag{C.7}$$

We can now define the α-twisted equivariant Hecke operators acting on the space of $SL(2,\mathbb{Z}) \times G$-equivariant Jacobi forms $\phi_{g,h}$ by

$$\mathcal{T}_L^\alpha \phi_{g,h}(\tau,z) := \frac{1}{L} \sum_{u \in SL(2,\mathbb{Z}) \backslash \mathrm{Mat}_L(\mathbb{Z})} \phi_{g,h}(\tau,z)|_\alpha(u,e) \ , \tag{C.8}$$

$$\mathcal{U}_L^\alpha \phi_{g,h}(\tau,z) := \phi_{g,h}(\tau,z)|_\alpha(\begin{smallmatrix} L & 0 \\ 0 & L \end{smallmatrix}, e) \ . \tag{C.9}$$

One may check that these operators satisfy the Hecke algebra (3.13)–(3.16). Furthermore, $\mathcal{T}_L^\alpha$ (respectively, $\mathcal{U}_L^\alpha$) maps the system of α-twisted $SL(2,\mathbb{Z})\times G$-equivariant (weak) Jacobi forms of weight 0 and index m to a system of α^L-twisted (respectively, α^{L^2}-twisted) $SL(2,\mathbb{Z})\times G$-equivariant (weak) Jacobi forms of weight 0 and index Lm (respectively, index L^2m). The proofs of these properties are completely analogous to the case where α is trivial; in fact, these properties follow directly from the definition of the Hecke operators in terms of slash operators satisfying a composition law of the form (C.7).

C.2 Definition of slash operator for $\mathrm{Mat}(\mathbb{Z})\times G$

In this subsection, we will describe the extension of the α-twisted slash operators to $\mathrm{Mat}(\mathbb{Z})\times G$. As stressed in section 3.2, it is sufficient to define such operators for matrices of the form $\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right)$, since, together with $SL(2,\mathbb{Z})$ they generate the whole $\mathrm{Mat}(\mathbb{Z})$. We will make an ansatz for the phase $\epsilon_{g,h}\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big)$ and verify that it extends consistently to the whole $\mathrm{Mat}(\mathbb{Z})$.

Our ansatz for the phase $\epsilon_{g,h}\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big)$ is based on the interpretation of the corresponding slash operator within the theory of symmetric orbifolds of conformal field theories. We recall that the twisted twining partition function $Z_{g,h}(\tau)$ in a holomorphic CFT is defined as a trace $\mathrm{Tr}_{\mathcal{H}_g}(\rho_g(h)q^{L_0-\frac{c}{24}})$, where $\mathcal{H}_g$ is the g-twisted sector and ρ_g is the (possibly projective) representation of the centralizer $C_G(g)$ on $\mathcal{H}_g$. This partition function can be computed by a path integral on a torus $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ where the fields are required to have monodromies g and h along the cycles -1 and τ, respectively. The function

$$Z_{g,h}(\tau)|_\alpha\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big) = \Upsilon^*_{\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right)} Z_{g,h}\ , \tag{C.10}$$

is associated with the L-fold covering $\mathbb{C}/(\mathbb{Z}+L\tau\mathbb{Z})$ and has a natural interpretation as a trace

$$Z_{g,h}(\tau)|_\alpha\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big) = \mathrm{Tr}_{\mathcal{H}_g}\big((\rho_g(h)q^{L_0-\frac{c}{24}})^L\big)\ . \tag{C.11}$$

Using the product law (2.5) for the projective representation ρ_g, we obtain

$$Z_{g,h}(\tau)|_\alpha\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big) = \prod_{i=1}^{L-1} c_g(h,h^i)\,\mathrm{Tr}_{\mathcal{H}_g}(\rho_g(h^L)q^{LL_0-\frac{Lc}{24}}) = \prod_{i=1}^{L-1} c_g(h,h^i) Z_{g,h^L}(L\tau)\ . \tag{C.12}$$

This formula suggests the definition

$$\epsilon_{g,h}\big(\left(\begin{smallmatrix} L & 0\\ 0 & 1\end{smallmatrix}\right), e\big) = \prod_{i=1}^{L-1} c_g(h,h^i)\ , \tag{C.13}$$

so that

$$f(g,h;\tau)|_\alpha\big(\left(\begin{smallmatrix} L & 0 \\ 0 & 1\end{smallmatrix}\right),e\big)=\prod_{i=1}^{L-1}c_g(h,h^i)f\Big((g,h)\left(\begin{smallmatrix} 1 & 0 \\ 0 & L\end{smallmatrix}\right);\,\left(\begin{smallmatrix} L & 0 \\ 0 & 1\end{smallmatrix}\right)\cdot\tau\Big)\ , \tag{C.14}$$

Any integer matrix of determinant L can be written as a product of elements in $SL(2,\mathbb{Z})$ and matrices of the form $\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)$ for prime p. For each $(u,k)\in\mathrm{Mat}(\mathbb{Z})\times G$, given a representation of u as a word $u=a_1\cdots a_n$ in the generators $a_1,\ldots,a_n$ of $\mathrm{Mat}(\mathbb{Z})$, one can use the composition law

$$f(g,h;\tau)|_\alpha(u_1u_2,k_1k_2)=f(g,h;\tau)|_\alpha(u_1,k_1)\ |_{\alpha^{|u_1|}}(u_2,k_2)\ , \tag{C.15}$$

where $|u|:=\det u$, to define the slash operator

$$f(g,h;\tau)|_\alpha(a_1\cdots a_n,k):=f(g,h;\tau)|_\alpha(\mathbb{I},k)\ |_\alpha(a_1,e)\ |_{\alpha^{|a_1|}}(a_2,e)\ \cdots\ |_{\alpha^{|a_1\cdots a_{n-1}|}}(a_n,e)\ . \tag{C.16}$$

One needs to check that this definition is consistent, i.e. that the operator $|_\alpha(u,k)$ does not depend on representation of $(u,k)\in\mathrm{Mat}(\mathbb{Z})\times G$ as a word in the generators and that the composition law is respected. The outline of the proof is given in the next subsection.

In particular, by (C.15), the identity

$$\begin{pmatrix}1 & 0\\ 0 & L\end{pmatrix}=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}L & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \tag{C.17}$$

yields

$$\epsilon_{g,h}\big(\left(\begin{smallmatrix} 1 & 0 \\ 0 & L\end{smallmatrix}\right),e\big)=\frac{c_g(h^{-1},h)^L\prod_{i=1}^{L-1}c_{h^{-1}}(g,g^i)}{c_{g^L}(h^{-1},h)}=\frac{1}{\prod_{i=1}^{L-1}c_h(g,g^i)}\ , \tag{C.18}$$

where in the last step we use (A.19), while from the identity

$$\begin{pmatrix}a & b\\ 0 & d\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & d\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^b\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix} \tag{C.19}$$

we obtain the general formula

$$\epsilon_{g,h}\left(\begin{smallmatrix} a & b \\ 0 & d\end{smallmatrix}\right):=\epsilon_{g,h}\big(\left(\begin{smallmatrix} a & b \\ 0 & d\end{smallmatrix}\right),e\big)=\frac{\prod_{i=1}^{a-1}c_g(h,h^i)^d}{\prod_{j=1}^{d-1}c_{g^{-b}h^a}(g,g^j)\prod_{k=1}^{b}c_g(g,g^{-k}h^a)^d}\ . \tag{C.20}$$

Since for any coset of $SL(2,\mathbb{Z})\backslash\,\mathrm{Mat}(\mathbb{Z})$ we can choose a representative of the form $\left(\begin{smallmatrix} a & b \\ 0 & d\end{smallmatrix}\right)$, this formula is sufficient to determine the Hecke operators $\mathcal{T}_L^\alpha$ and $\mathcal{U}_L^\alpha$ for all L. In particular, if the restriction of the 3-cocycle α to the group $\langle g,h\rangle\cong\mathbb{Z}_{N_1}\times\mathbb{Z}_{N_2}$ is given by $\alpha=\alpha_{v_1}\alpha_{v_2}\alpha_{v_{12}}$, in terms of the generators (A.20)–(A.21), then $(N_1,N_2>1)$

$$\epsilon_{g,h}\big(\left(\begin{smallmatrix} a & b \\ 0 & d\end{smallmatrix}\right),e\big)=e^{\frac{2\pi i}{N_1}\left(v_1b\sum_{j=1}^{d-1}\delta^{(N_1)}(j+1)-v_1d\sum_{j=1}^{b}\delta^{(N_1)}(1-j)+v_{12}d\sum_{j=1}^{a-1}\delta^{(N_2)}(j+1)\right)}\ . \tag{C.21}$$

C.3 Proof of consistency

In the rest of this section, we will show that our proposal for the twisted equivariant Hecke operators satisfies several consistency conditions. As a first consistency, let us consider the effect of a 'gauge transformation' (A.5), under which the normalized 3-cocycle α is multiplied by a 3-coboundary $\partial\beta$ (with $\beta(e,g)=\beta(g,e)=1$ to keep the normalization). Formally, this transformation corresponds to a different choice of trivialization of the pull-back $\pi^*\mathcal{L}^\alpha_{g,h}$, where $\pi:\mathcal{P}\times\mathbb{H}_+\times\mathbb{C}\to\mathcal{M}$ is the covering map. As stressed in the previous appendix A, for a holomorphic CFTs $\mathcal{C}$, the transformation (A.5) corresponds to a redefinition (A.15) of the projective representations ρ_g of the centralizer $C_G(g)$ over the twisted sector $\mathcal{H}_g$ and induces the transformation (A.16) of the partition functions. For the tensor product $\mathcal{C}^{\otimes L}$ and the symmetric product $S^L\mathcal{C}$ theories, the twisted twining partition functions $Z_{g,h}$ are defined as traces over the representation induced by the (symmetrized) L-tensor product of ρ_g, so that the appropriate transformation is

$$Z^{(L)}_{g,h}\mapsto\tilde{Z}^{(L)}_{g,h}:=\gamma_g(h)^L Z^{(L)}_{g,h}\ , \tag{C.22}$$

where $\gamma_g(h)=\beta(g,h)/\beta(h,g)$. Analogous properties hold for the twisted twining genera $\phi^{(L)}_{g,h}$ in superconformal field theories. Using the identities $\gamma_e(g)=\gamma_g(e)=1$ and $\gamma_g(h)=\gamma_h(g)^{-1}$, it is easy to check that, under (A.15) and (A.11),

$$\big(\gamma_g(h)f(g,h;\tau)\big)|_{\tilde\alpha}\big(\left(\begin{smallmatrix}a&b\\0&d\end{smallmatrix}\right)\gamma,e\big)=\gamma_g(h)^L\Big(f(g,h;\tau)|_\alpha\big(\left(\begin{smallmatrix}a&b\\0&d\end{smallmatrix}\right)\gamma,e\big)\Big)\ , \tag{C.23}$$

where $L=ad$, and similar properties hold for Jacobi forms. As a consequence, we obtain

$$\mathcal{T}^{\tilde\alpha}_L\tilde{Z}_{g,h}=\gamma_g(h)^L\mathcal{T}^\alpha_L Z_{g,h}\ , \tag{C.24}$$

and

$$\mathcal{T}^{\tilde\alpha}_L\tilde\phi_{g,h}=\gamma_g(h)^L\mathcal{T}^\alpha_L\phi_{g,h}\ ,\qquad \mathcal{U}^{\tilde\alpha}_L\tilde\phi_{g,h}=\gamma_g(h)^{L^2}\mathcal{U}^\alpha_L\phi_{g,h}\ , \tag{C.25}$$

which indeed reproduce (C.22) and the analogous formula for twisted twining genera. More generally, given *any* representation of a matrix $u\in\mathrm{Mat}_L(\mathbb{Z})$ as a word in the generators of $\mathrm{Mat}(\mathbb{Z})$, it is easy to check that (C.23) holds for the slash operator $|_\alpha(u,k)$, i.e

$$\gamma_g(h)\tilde\epsilon_{g,h}(u,k)=\gamma_g(h)^L\epsilon_{g,h}(u,k)\ . \tag{C.26}$$

An easy corollary of this property is

$$f(g,h;\tau)|_\alpha(u,k)|_{\alpha^L}(\mathbb{I},x)=f(g,h;\tau)|_\alpha(\mathbb{I},x)|_\alpha(u,k)\ ,\qquad (u,k)\in\mathrm{Mat}_L(\mathbb{Z})\times G,\ x\in G\ , \tag{C.27}$$

which follows by considering (C.26) for the transformation (A.7). As a consequence of (C.27), one only needs to prove consistency of the slash operator for elements of the form (u, e).

A set of generators of the multiplicative semigroup $\mathrm{Mat}(\mathbb{Z})$ is given by the generators S, T of $SL(2,\mathbb{Z})$ (see eq.(B.9)), together with all matrices $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for p prime. To check the consistency of the definition of the slash operator, one needs to verify that, for each relation $a_1 \cdots a_n = b_1 \cdots b_m$ among the generators of $\mathrm{Mat}(\mathbb{Z})$, the following identities hold for all $(g, h) \in \mathcal{P}_G$

$$f(g,h;\tau)|_\alpha (a_1, e) \cdots |_{\alpha^{|a_1 \cdots a_{n-1}|}} (a_n, e) = f(g,h;\tau)|_\alpha (b_1, e) \cdots |_{\alpha^{|b_1 \cdots b_{m-1}|}} (b_m, e) \ . \tag{C.28}$$

This is equivalent to checking the following identities for the phases

$$\prod_{i=1}^{n} \epsilon_{(g,h)a_{i+1}^\vee \cdots a_n^\vee}(a_i, e)^{|a_1 \cdots a_{i-1}|} = \prod_{i=1}^{m} \epsilon_{(g,h)b_{i+1}^\vee \cdots b_m^\vee}(b_i, e)^{|b_1 \cdots b_{i-1}|} \ . \tag{C.29}$$

For the relations $S^4 = 1$ and $(ST)^3 = S^2$ within $SL(2,\mathbb{Z})$, these conditions follow from the identities in [46, 47]; therefore, we will only consider the independent relations within $\mathrm{Mat}_L(\mathbb{Z})$ for $L > 1$.

For each prime power p, the relations within $\mathrm{Mat}_p(\mathbb{Z})$ can be obtained from the relations for $SL(2,\mathbb{Z})$, together with relations of the form

$$\gamma_1 \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \gamma_2 \ , \tag{C.30}$$

for suitable $\gamma_1, \gamma_2 \in SL(2,\mathbb{Z})$. It is easy to check that such a relation exists if and only if γ_2 is an element of $\Gamma_0(p)$, which is generated by T, S^2 and $ST^pS = \left(\begin{smallmatrix} 1 & 0 \\ -p & 1 \end{smallmatrix}\right)$. Therefore, the only independent relations within $\mathrm{Mat}_p(\mathbb{Z})$ are

$$T^p \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} T \ , \tag{C.31}$$

$$S^2 \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} S^2 \ , \tag{C.32}$$

and

$$STS \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} ST^pS \ . \tag{C.33}$$

The only remaining relations in $\mathrm{Mat}(\mathbb{Z})$ are of the form

$$\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \gamma_1 \begin{pmatrix} p' & 0 \\ 0 & 1 \end{pmatrix} = \gamma_2 \begin{pmatrix} p' & 0 \\ 0 & 1 \end{pmatrix} \gamma_3 \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \gamma_4 \ , \tag{C.34}$$

for all pairs of distinct primes p, p' and suitable $\gamma_1, \gamma_2, \gamma_3, \gamma_4 \in SL(2,\mathbb{Z})$. In fact, it is easier to consider formula (C.20) as a definition of $\epsilon_{g,h}((\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}), e)$ and verify the identities (C.29) for the relations

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} pa & b \\ 0 & d \end{pmatrix} , \tag{C.35}$$

$$\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} pa & pb \\ 0 & d \end{pmatrix} , \tag{C.36}$$

since the consistency conditions for the relations (C.34) then follow. The consistency conditions are therefore

$$\epsilon_{g,h^p}\big(\begin{smallmatrix} 1 & p \\ 0 & 1 \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big) = \epsilon_{g,g^{-1}h}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\big)^p \tag{C.37}$$

$$\epsilon_{h^{-p},g}\big(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\big)\epsilon_{g,h^p}\big(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big) = \epsilon_{g^{-1},h^{-1}}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big)\epsilon_{h^{-1},p}\big(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\big)^p\epsilon_{g,h}\big(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\big)^p \tag{C.38}$$

$$\epsilon_{g^p,h}\big(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\big) = \epsilon_{g,g^{-p}h}\big(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} 1 & p \\ 0 & 1 \end{smallmatrix}\big)^p \tag{C.39}$$

$$\epsilon_{g,h^p}\big(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big)^{ad} = \epsilon_{g,h}\big(\begin{smallmatrix} pa & b \\ 0 & d \end{smallmatrix}\big) \tag{C.40}$$

$$\epsilon_{g^d,g^{-b}h^a}\big(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\big)\epsilon_{g,h}\big(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\big)^p = \epsilon_{g,h}\big(\begin{smallmatrix} pa & pb \\ 0 & d \end{smallmatrix}\big) . \tag{C.41}$$

The proof is a tedious but straightforward computation, consisting of a repeated use of the cocycle conditions. For example, the proof of (C.40):

$$\begin{aligned} &\frac{\prod_{i=1}^{a-1} c_g(h,h^{ip})^d}{\prod_{j=1}^{d-1} c_{g^{-b}h^{ap}}(g,g^j)\prod_{k=1}^{b} c_g(g,g^{-k}h^{ap})^d} \prod_{i=1}^{p-1} c_g(h,h^i)^{ad} \frac{\prod_{j=1}^{d-1} c_{g^{-b}h^{ap}}(g,g^j)\prod_{k=1}^{b} c_g(g,g^{-k}h^{ap})^d}{\prod_{i=1}^{pa-1} c_g(h,h^i)^d} \\ &= \Big(\frac{\prod_{i=1}^{a-1} c_g(h,h^{ip})\prod_{i=1}^{p-1} c_g(h,h^i)^a}{\prod_{i=1}^{pa-1} c_g(h,h^i)}\Big)^d = 1 \end{aligned} \tag{C.42}$$

where the last equality follows from the 2-cocycle condition on c_g.

D Some technical proofs

D.1 Proof of (3.47).

In this section, we will prove eq.(3.47). The sum $F_{g,h}(a,d,k,\ell)$ can be written as

$$F_{g,h}(a,d,k,\ell) := \sum_{n=0}^{\infty} \frac{1}{d} \sum_{b=0}^{d-1} e^{\frac{2\pi i b n}{N\lambda d}} c_{g_\alpha^d, g_\alpha^{-b} h_\alpha^k}\big(\frac{n}{N\lambda}, \ell\big) q^{\frac{na}{N\lambda d}} , \tag{D.1}$$

in terms of the Fourier coefficients of the twisted twining genera $\phi_{g_\alpha^d, g_\alpha^{-b} h_\alpha^k}$ introduced in section 3.2.3. Notice that the functions $\phi_{g_\alpha, h_\alpha}$ satisfy

$$\phi_{g_\alpha^r, g_\alpha^{-r} h_\alpha}(\tau+1, z) = \phi_{g_\alpha^r, h_\alpha}(\tau, z) \ , \tag{D.2}$$

so that

$$c_{g_\alpha^d, g_\alpha^d h_\alpha}(r, \ell) = e^{\frac{2\pi i r}{N\lambda}} c_{g_\alpha^d, h_\alpha}(r, \ell) \ , \tag{D.3}$$

for all $h_\alpha \in C^\alpha_{M_{24}}(g)$. Set

$$e := \gcd(d, N\lambda) \ , \qquad f := \frac{d}{e} \ , \tag{D.4}$$

and note that

$$\gcd(f, \frac{N\lambda}{e}) = 1 \ , \tag{D.5}$$

otherwise $e \gcd(f, \frac{N\lambda}{e})$ would be a common divisor of d and $N\lambda$ greater than e. The order of g_α^d is $o(g_\alpha^{ef}) = o(g_\alpha^e) = N\lambda/e$, so that $c_{g_\alpha^d, g_\alpha^{-b} h_\alpha^k}(\frac{n}{N\lambda}, \ell) = 0$ unless $e|n$. Thus, we can set $r := n/e$ and obtain

$$F_{g,h}(a, d, k, \ell) = \sum_{r=0}^{\infty} \frac{1}{d} \sum_{b=0}^{d-1} e^{\frac{2\pi i b r e}{N\lambda d}} c_{g_\alpha^d, g_\alpha^{-b} h_\alpha^k}(\frac{re}{N\lambda}, \ell) q^{\frac{ar}{N\lambda f}} \ . \tag{D.6}$$

Notice that, by (D.3), the general term of the sum over b is periodic under $b \to b+d$. We can set $b := se + b'$ and replace the sum over $b \in \mathbb{Z}/d\mathbb{Z}$ by a sum over $b' \in \mathbb{Z}/e\mathbb{Z}$ and $s \in \mathbb{Z}/f\mathbb{Z}$

$$F_{g,h}(a, d, k, \ell) = \sum_{r=0}^{\infty} \frac{1}{e} \sum_{b'=0}^{e-1} \frac{1}{f} \sum_{s=0}^{f-1} e^{\frac{2\pi i (b'+se) r}{N\lambda f}} c_{g_\alpha^d, g_\alpha^{-b'} g_\alpha^{-se} h_\alpha^k}(\frac{re}{N\lambda}, \ell) q^{\frac{ar}{N\lambda f}} \ . \tag{D.7}$$

Thus, the general term of the sum over b' is periodic under $b' \to b' + e$ and we can sum over $N\lambda/e$ periods and divide by $N\lambda/e$

$$F_{g,h}(a, d, k, \ell) = \sum_{r=0}^{\infty} \frac{1}{N\lambda} \sum_{b'=0}^{N\lambda-1} \frac{1}{f} \sum_{s=0}^{f-1} e^{\frac{2\pi i (b'+se) r}{N\lambda f}} c_{g_\alpha^d, g_\alpha^{-b'} g_\alpha^{-se} h_\alpha^k}(\frac{re}{N\lambda}, \ell) q^{\frac{ar}{N\lambda f}} \ . \tag{D.8}$$

By (D.5), there are integers x, y such that

$$xf + y\frac{N\lambda}{e} = 1 \tag{D.9}$$

so that $se = sxd + syN\lambda$ and we obtain

$$\begin{aligned} F_{g,h}(a, d, k, \ell) =& \sum_{r=0}^{\infty} \frac{1}{N\lambda} \sum_{b'=0}^{N\lambda-1} \frac{1}{f} \sum_{s=0}^{f-1} e^{\frac{2\pi i b' r}{N\lambda f}} e^{\frac{2\pi i s (xd + yN\lambda) r}{N\lambda f}} c_{g_\alpha^d, g_\alpha^{-b'} g_\alpha^{-sdx} h_\alpha^k}(\frac{re}{N\lambda}, \ell) q^{\frac{ar}{N\lambda f}} \\ =& \sum_{r=0}^{\infty} \sum_{b'=0}^{N\lambda-1} \frac{e^{\frac{2\pi i b' r}{N\lambda f}}}{N\lambda} \sum_{s=0}^{f-1} \frac{e^{\frac{2\pi i s y r}{f}}}{f} c_{g_\alpha^d, g_\alpha^{-b'} h_\alpha^k}(\frac{re}{N\lambda}, \ell) q^{\frac{ar}{N\lambda f}} \ , \end{aligned} \tag{D.10}$$

where in the last step we used (D.3). Since by (D.9) $\gcd(y,f)=1$, the sum over s just a projection

$$\sum_{s=0}^{f-1}\frac{e^{\frac{2\pi i s y r}{f}}}{f}=\begin{cases}1 & \text{if } f|r\ ,\\ 0 & \text{otherwise}\ ,\end{cases} \tag{D.11}$$

so that, by setting $r=mf$, we obtain

$$F_{g,h}(a,d,k,\ell)=\sum_{m=0}^{\infty}\sum_{b'=0}^{N\lambda-1}\frac{e^{\frac{2\pi i b' m}{N\lambda}}}{N\lambda}c_{g_\alpha^d,g_\alpha^{-b'}h_\alpha^k}(\frac{md}{N\lambda},\ell)q^{\frac{am}{N\lambda}}\ , \tag{D.12}$$

that is equivalent to (3.47).

D.2 Central extensions and special choices of the cocycle

In section 3.2.3 a central extension $C^\alpha_{M_{24}}(g)$ of the centralizer $C_{M_{24}}(g)$ is defined, together with the representations $\tilde\rho_{g,r}$ on the g^r-twisted sector $\mathcal{H}_{g^r}$, for all $r\in\mathbb{Z}_{\geq 0}$. In this section, we will describe some special choices of the cocycle α, for which these representations are particularly simple.

If two 3-cocycles α, α' differ by a coboundary $\partial\beta$, then there is an isomorphism $C^{\alpha'}_{M_{24}}(g)\xrightarrow{\cong}C^\alpha_{M_{24}}(g)$ that relates the canonical lifts

$$h_{\alpha'}\mapsto h_\alpha\, q(\nu_g(h))\ , \tag{D.13}$$

where $\nu_g(h)\in\mathbb{R}/\mathbb{Z}$ is defined by $e^{2\pi i\nu_g(h)}=\frac{\beta(g,h)}{\beta(h,g)}$.

Under the shift $\alpha\to\alpha'$ by a coboundary $\partial\beta$, the cocycle c_g transforms as

$$c_g(h,k)\to\frac{\beta(g,h)}{\beta(h,g)}\frac{\beta(g,k)}{\beta(k,g)}\frac{\beta(hk,g)}{\beta(g,hk)}c_g(h,k)\ ,\qquad h,k\in C_{M_{24}}(g)\ . \tag{D.14}$$

Correspondingly, the phases $f_{g,r}(h)$ of eq.(3.31) transform as

$$f_{g,r}(h)\to f_{g,r}(h)\frac{\beta(h,g)^r}{\beta(g,h)^r}\frac{\beta(g^r,h)}{\beta(h,g^r)}\ ,\qquad h\in C_{M_{24}}(g)\ , \tag{D.15}$$

and, in particular,

$$f_{g,N}(h)\to f_{g,N}(h)\frac{\beta(h,g)^N}{\beta(g,h)^N}\ ,\qquad h\in C_{M_{24}}(g)\ . \tag{D.16}$$

Therefore, we can choose $\beta(g,h)/\beta(h,g)$ for each $h\in C_{M_{24}}(g)$, $h\notin\langle g\rangle$, in such a way that

$$f_{g,N}(h)=1\ ,\qquad \forall h\in C_{M_{24}}(g)\setminus\langle g\rangle\ . \tag{D.17}$$

This condition determines $\beta(g,h)/\beta(h,g)$ for all $h \in C_{M_{24}}(g) \setminus \langle g \rangle$ up to N-th roots of unity. Notice that, in general,

$$\tilde{\rho}_{g,r+N}(h) = \frac{\rho_{g^r}(h)}{f_{g,r}(h) f_{g,N}(h)} = \frac{\tilde{\rho}_{g^r}(h)}{f_{g,N}(h)} \ , \qquad h \in C_{M_{24}}(g) \ , \tag{D.18}$$

so that, by imposing (D.17), we have

$$\tilde{\rho}_{g,r+N}(h) = \tilde{\rho}_{g^r}(h) \ , \qquad h \in C_{M_{24}}(g) \setminus \langle g \rangle \ . \tag{D.19}$$

On the other hand, $f_{g,N}(g)$ depends only on the cohomology class $[\alpha]$. Since the restriction of $[\alpha]^N$ to $H^3(\langle g \rangle, U(1)) \cong \mathbb{Z}_N$ is the trivial class, it follows that $f_{g,N}(g)$ must be a N-th root of unity. In fact, for M_{24}, we have

$$f_{g,N}(g) = e^{-\frac{2\pi i}{\lambda}} \ , \tag{D.20}$$

where $\lambda | N$ is the length of the shortest cycle of g in the 24-dimensional permutation representation. It is also related to the spectrum of $L_0 - \frac{c}{24}$ in the g-twisted sector, which takes values in $-\frac{1}{\lambda N} + \frac{1}{N}\mathbb{Z}$ and to the presence of a non-trivial multiplier system for the twining genus $\phi_{e,g}$. It is useful to distinguishes the cases when $\lambda = 1$ and $\lambda \neq 1$.

D.2.1 Case $\lambda = 1$ (trivial multiplier)

This case occurs whenever g is an element of some M_{23} subgroup of M_{24}, i.e. when g belongs to the classes

$$2A,\ 3A,\ 4B,\ 5A,\ 6A,\ 7A,\ 7B,\ 8A,\ 11A,\ 14A,\ 14B,\ 15A,\ 15B,\ 23A,\ 23B \ .$$

In this case the restriction of $[\alpha]$ to $H^3(\langle g \rangle, U(1))$ is the trivial class and one can choose $\beta(g^i, g^j)$ in such a way that

$$c_{g^i}(g^j, g^k) = 1 \ . \tag{D.21}$$

In particular,

$$f_{g,r}(g^i) = 1 \ , \tag{D.22}$$

for all r, i. By specializing (3.32) to the case $r = N$, we obtain

$$c_g(h,k)^N = c_{g^N}(h,k) \frac{f_{g,N}(hk)}{f_{g,N}(h) f_{g,N}(k)} = 1 \ , \qquad h,k \in C(g) \ , \tag{D.23}$$

so that $c_g(h,k)$ is an N-th root of unity for all $h,k \in C_{M_{24}}(g)$, i.e. there is $\mu_g(h,k) \in \mathbb{Z}/N\mathbb{Z}$ such that

$$c_g(h,k) = e^{\frac{2\pi i \mu_g(h,k)}{N}} \ . \tag{D.24}$$

With this choice of cocycle, the central extension $C^{\alpha}_{M_{24}}(g)$ of section 3.2.3 can be chosen to be finite

$$1 \to \langle Q \rangle \cong \mathbb{Z}_N \to C^{\alpha}_{M_{24}}(g) \to C_{M_{24}}(g) \to 1 \ . \tag{D.25}$$

where the central element Q is related to the $U(1)$-generator $q(x)$ in section 3.2.3 by

$$Q = q(1/N) \ . \tag{D.26}$$

D.2.2 Case $\lambda \neq 1$ (non-trivial multiplier)

This case occurs when g is in one of the M_{24}-classes

$$2B,\ 3B,\ 4A,\ 4C,\ 6B,\ 10A,\ 12A,\ 12B,\ 21A,\ 21B \ .$$

Since $H^3(\langle g \rangle, U(1)) \cong \mathbb{Z}_N$, we can choose the cocycle in such a way that

$$c_{g^i}(g^j, g^k)^N = 1 \ , \tag{D.27}$$

for all $i, j, k \in \mathbb{Z}$. This condition fixes $\beta(g^i, g^j)$ up to N-th roots of unity. For $h, k \in C_{M_{24}}(g)$, with $h, k, hk \notin \langle g \rangle$, we can apply again (3.32) with $r = N$ and obtain

$$c_g(h, k)^N = 1 \ , \qquad h, k, hk \in C_{M_{24}}(g) \setminus \langle g \rangle \ . \tag{D.28}$$

If α satisfies (D.17) and (D.27), then by (3.32) we have

$$f_{g,N}(g^i) c_g(g^i, h)^N = 1 \ , \qquad \frac{f_{g,N}(g^i) f_{g,N}(g^j)}{f_{g,N}(g^{i+j})} = 1 \ , \tag{D.29}$$

so that

$$f_{g,N}(g^n) = f_{g,N}(g)^n = e^{-\frac{2\pi i n}{\lambda}} \ , \tag{D.30}$$

and

$$c_g(g^i, h)^{N\lambda} = 1 \ . \tag{D.31}$$

Therefore, $c_g(h, k)$ is a $N\lambda$ roots of unity for all $h, k \in C_{M_{24}}(g)$ and we can define $\mu_g(h, k) \in \mathbb{Z}/N\lambda\mathbb{Z}$ such that

$$c_g(h, k) = e^{\frac{2\pi i \mu_g(h,k)}{N\lambda}} \ . \tag{D.32}$$

The discussion is similar to the case $\lambda = 1$. With this choice of cocycle, the central extension $C^{\alpha}_{M_{24}}(g)$ of section 3.2.3 can be chosen to be finite

$$1 \to \langle q(\frac{1}{N\lambda}) \rangle \cong \mathbb{Z}_{N\lambda} \to C^{\alpha}_{M_{24}}(g) \to C_{M_{24}}(g) \to 1 \ . \tag{D.33}$$

It is convenient to define also a central element Q of order N by

$$Q = q(1/N) \ , \tag{D.34}$$

so that the g^r-twisted sector is an eigenspace of Q with eigenvalue $e^{\frac{2\pi i r}{N}}$.

D.3 S-duality for g in classes $2B, 3B, 4C, 6B, 12B$

Here we prove the S-duality transformation property of $\Phi_{g,h}$ for g in the classes 2B, 3B, 4C, 6B, 12B. Let g be an element of M_{24} in one of these classes and let N be its order. In all these cases, the restriction of the class $[\alpha]$ to $H^3(\langle g\rangle, U(1))$ is non-trivial and has order $\lambda = N$. Therefore,

$$\hat{c}_{g,h}(d,m,\ell,t) = \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b m}{N^2}}}{N^2} \operatorname{Tr}_{\mathcal{H}_{g^d}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,d}(g)^{-b}\tilde{\rho}_{g,d}(h)^k(-1)^{F+\bar{F}}\big) . \tag{D.35}$$

In the g^r-twisted sector, for r and N coprime, the spectrum of $L_0 - \frac{c}{24}$ takes values in $-\frac{1}{N^2} + \frac{1}{N}\mathbb{Z}$. More generally, if $\gcd(r,N) = e$, then in the g^r-twisted sector the spectrum of $L_0 - \frac{c}{24}$ takes values in $-\frac{1}{(N/e)^2} + \frac{1}{N/e}\mathbb{Z}$. Notice that $\hat{c}_{g,h}(d,m,\ell,t)$ is a sum of traces in the g^d-twisted sector at level $\frac{md}{N^2}$, so that $\hat{c}_{g,h}(d,m,\ell,t) = 0$ unless $\frac{md}{N^2} \in -\frac{1}{(N/e)^2} + \frac{1}{N/e}\mathbb{Z}$, where $e = \gcd(N,d)$. It is easy to verify that this condition is equivalent to

$$m \equiv -d \mod N , \tag{D.36}$$

where we used the property

$$\gcd(x,N) = 1 \qquad \Leftrightarrow \qquad x^2 \equiv 1 \mod N , \tag{D.37}$$

which holds for all N that divide 24.

As explained in appendix D.2.2, we can choose the cocycle α in such a way that

$$f_{g,N}(h) \equiv \prod_{i=1}^{N-1} c_h(g,g^i) = 1 , \qquad h \in C_{M_{24}}(g) \setminus \langle g\rangle , \tag{D.38}$$

while $f_{g,N}(g^k)$ depends only on the cohomology class $[\alpha]$ and equals

$$f_{g,N}(g^k) = e^{-\frac{2\pi i k}{N}} . \tag{D.39}$$

With this choice of cocycle, we have

$$\tilde{\rho}_{g,x+N}(h) = \frac{\tilde{\rho}_{g,x}(h)}{f_{g,N}(h)} = \tilde{\rho}_{g,x}(h) , \qquad h \in C_{M_{24}}(g) \setminus \langle g\rangle , \tag{D.40}$$

while

$$\tilde{\rho}_{g,x+N}(g) = \frac{\tilde{\rho}_{g,x}(g)}{f_{g,N}(g)} = e^{\frac{2\pi i}{N}}\tilde{\rho}_{g,x}(g) . \tag{D.41}$$

For $d \equiv -m \equiv 0 \mod N$, we have, for all $x \in \mathbb{N}$,

$$\begin{aligned}\hat{c}_{g,h}(d = Nu, m = Nv, \ell, t) &= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b v}{N}}}{N^2} \operatorname{Tr}_{\mathcal{H}_e(uv,\ell)}\big(\tilde{\rho}_{g,Nu}(g)^{-b} \tilde{\rho}_{g,Nu}(h)^k (-1)^{F+\bar{F}}\big) \\ &= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(v-u)}{N}}}{N^2} \operatorname{Tr}_{\mathcal{H}_e(uv,\ell)}\big(g^{-b} h^k (-1)^{F+\bar{F}}\big) \ . \end{aligned} \tag{D.42}$$

Since the (untwisted) twining genera are invariant under charge conjugation, i.e.

$$\phi_{e,g^{-b}h^k}(\tau, z) = \phi_{e,g^b h^{-k}}(\tau, z) \ , \tag{D.43}$$

we obtain

$$\begin{aligned}\hat{c}_{g,h}(Nu, Nv, \ell, t) &= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(v-u)}{N}}}{N^2} \operatorname{Tr}_{\mathcal{H}_e(uv,\ell)}\big(g^b h^{-k} (-1)^{F+\bar{F}}\big) \\ = \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} &\frac{e^{\frac{2\pi i b(u-v)}{N}}}{N^2} \operatorname{Tr}_{\mathcal{H}_e(uv,\ell)}\big(g^{-b} h^{-k} (-1)^{F+\bar{F}}\big) = \hat{c}_{g,h^{-1}}(Nv, Nu, \ell, t) \ . \end{aligned} \tag{D.44}$$

Similarly, for $d \equiv x \equiv -m \mod N$, with $x = 1, \ldots, N-1$, we obtain

$$\begin{aligned}\hat{c}_{g,h}(d = Nu + x, m = Nv + N - x, \ell, t) &= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(Nv+N-x-Nu)}{N^2}}}{N^2} \\ &\quad \times \operatorname{Tr}_{\mathcal{H}_{g^x}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,x}(g)^{-b} \tilde{\rho}_{g,x}(h)^k (-1)^{F+\bar{F}}\big) \ . \end{aligned} \tag{D.45}$$

Using the relation

$$\phi_{g,h}(\tau, z) = \frac{1}{c_h(g, g^{-1}) c_{g^{-1}}(h, h^{-1})} \phi_{g^{-1},h^{-1}}(\tau, z) \ , \tag{D.46}$$

and by

$$\frac{\rho_{g^{-1}}(h^{-1})}{c_{g^{-1}}(h, h^{-1})} = \rho_{g^{-1}}(h)^{-1} \ , \tag{D.47}$$

we obtain

$$\begin{aligned}\operatorname{Tr}_{H_{g^x}(n,\ell)}(\tilde{\rho}_{g,x}(h)) &= \frac{\operatorname{Tr}_{\mathcal{H}_{g^x}(n,\ell)}(\rho_{g^x}(h))}{\prod_{i=1}^{x-1} c_h(g, g^i)} = \frac{\operatorname{Tr}_{\mathcal{H}_{g^{-x}}(n,\ell)}(\rho_{g^{-x}}(h)^{-1})}{c_h(g^x, g^{-x}) \prod_{i=1}^{x-1} c_h(g, g^i)} \\ = \frac{\operatorname{Tr}_{\mathcal{H}_{g^{N-x}}(n,\ell)}(\tilde{\rho}_{g,N-x}(h)^{-1})}{c_h(g^x, g^{N-x}) \prod_{i=1}^{x-1} c_h(g, g^i) \prod_{j=1}^{N-x-1} c_h(g, g^i)} &= \frac{\operatorname{Tr}_{\mathcal{H}_{g^{N-x}}(n,\ell)}(\tilde{\rho}_{g,N-x}(h)^{-1})}{f_{g,N}(h)} \ , \end{aligned} \tag{D.48}$$

where the last equality follows from

$$c_h(g^x, g^{N-x}) \prod_{i=1}^{x-1} c_h(g, g^i) = \prod_{i=N-x}^{N} c_h(g, g^i) \ , \tag{D.49}$$

that in turn is a consequence of the 2-cocycle condition for c_g. It is easy to verify that this identity is compatible with the product in the central extension, i.e.

$$\mathrm{Tr}_{H_{g^x}(n,\ell)}(\tilde{\rho}_{g,x}(h)\tilde{\rho}_{g,x}(k)) = \frac{\mathrm{Tr}_{\mathcal{H}_{g^{N-x}}(n,\ell)}(\tilde{\rho}_{g,N-x}(k)^{-1}\tilde{\rho}_{g,N-x}(h)^{-1})}{f_{g,N}(k) f_{g,N}(h)} \ . \tag{D.50}$$

Using this relation, we obtain

$$\begin{aligned}
&\hat{c}_{g,h}(d = Nu + x, m = Nv + N - x, \ell, t) \\
&= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(Nv-x-Nu)}{N^2}}}{N^2} \mathrm{Tr}_{\mathcal{H}_{g^{N-x}}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,N-x}(g)^b \tilde{\rho}_{g,N-x}(h)^{-k}(-1)^{F+\bar{F}}\big) \\
&= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(Nu+x-Nv)}{N^2}}}{N^2} \mathrm{Tr}_{\mathcal{H}_{g^{N-x}}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,N-x}(g)^{-b} \tilde{\rho}_{g,N-x}(h)^{-k}(-1)^{F+\bar{F}}\big) \\
&= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b(Nu+x)}{N^2}}}{N^2} \mathrm{Tr}_{\mathcal{H}_{g^{N-x}}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,Nv+N-x}(g)^{-b} \tilde{\rho}_{g,Nv+N-x}(h)^{-k}(-1)^{F+\bar{F}}\big) \\
&= \sum_{k=0}^{M-1} \frac{e^{-\frac{2\pi i t k}{M}}}{M} \sum_{b=1}^{N^2} \frac{e^{\frac{2\pi i b d}{N^2}}}{N^2} \mathrm{Tr}_{\mathcal{H}_{g^{m}}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,m}(g)^{-b} \tilde{\rho}_{g,m}(h)^{-k}(-1)^{F+\bar{F}}\big) \ .
\end{aligned} \tag{D.51}$$

Finally, we notice that

$$\tilde{\rho}_{g,r}(h)^{-k} = c_g(h, h^{-1})^{-rk} \tilde{\rho}_{g,r}(h^{-1})^k \ , \tag{D.52}$$

where $c_g(h, h^{-1})$ is a N-th root of unity thanks to our choice of cocycle. We recall that imposing the condition $f_{g,N}(h) = 1$ still leaves the possibility of modifying the cocycle α by a coboundary $\partial\beta$ with $\beta^N = 1$. Under such a modification, $c_g(h, h^{-1})$ transforms as

$$c_g(h, h^{-1}) \to \frac{\beta(g, h)}{\beta(h, g)} \frac{\beta(g, h^{-1})}{\beta(h^{-1}, g)} c_g(h, h^{-1}) \ , \tag{D.53}$$

We distinguish two cases. If h^{-1} and h are not conjugated within $C_{M_{24}}(g)$, then one can choose the cocycle α in such a way that

$$c_g(h, h^{-1}) = 1 \ . \tag{D.54}$$

With this choice, we obtain

$$\hat{c}_{g,h}(d, m, \ell, t) = \hat{c}_{g,h^{-1}}(m, d, \ell, t) \ , \tag{D.55}$$

so that

$$\begin{aligned}\Phi_{g,h}(\sigma,\tau,z) &= pq^{\frac{1}{N^2}}y\prod_{t=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N^2}}y^\ell p^d)^{\hat{c}_{g,h}(d,m,\ell,t)}\\ &= pq^{\frac{1}{N^2}}y\prod_{t=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N^2}}y^\ell p^d)^{\hat{c}_{g,h^{-1}}(m,d,\ell,t)}\\ &= \Phi_{g,h^{-1}}(\frac{\tau}{N^2},N^2\sigma,z)\ .\end{aligned} \tag{D.56}$$

In most cases, however, there is some $w \in C_{M_{24}}(g)$ such that $h^{-1} = w^{-1}hw$. At the level of the central extension $C^{\alpha}_{M_{24}}(g)$, the relation $hw = wh^{-1}$ leads to a relation among the lifts

$$h_\alpha^{-1} = \frac{c_g(w,h^{-1})}{c_g(h,w)c_g(h,h^{-1})}w_\alpha^{-1}h_\alpha w_\alpha = w_\alpha^{-1}Q^n h_\alpha w_\alpha\ , \tag{D.57}$$

where $n \in \mathbb{Z}/N\mathbb{Z}$ is such that

$$e^{-\frac{2\pi in}{N}} = \frac{c_g(h,w)c_g(h,h^{-1})}{c_g(w,h^{-1})}\ , \tag{D.58}$$

and we used the fact that, with our choice of cocycle, the right-hand side is a N-th root of unity. Since $Q^n h_\alpha$ is conjugated with h_α^{-1} within $C^{\alpha}_{M_{24}}(g)$, it must have the same order M, so that $nM \equiv 0 \mod N$ or, equivalently,

$$nM = sN\ , \tag{D.59}$$

for some $s \in \mathbb{Z}/M\mathbb{Z}$. Therefore,

$$\begin{aligned}&\hat{c}_{g,h}(d,m,\ell,t)\\ &\quad= \sum_{k=0}^{M-1}\frac{e^{-\frac{2\pi itk}{M}}}{M}\sum_{b=1}^{N^2}\frac{e^{\frac{2\pi ibd}{N^2}}}{N^2}\,\mathrm{Tr}_{\mathcal{H}_{g^m}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,m}(g)^{-b}\tilde{\rho}_{g,m}(h)^k\tilde{\rho}_{g,m}(Q^n)^k(-1)^{F+\bar{F}}\big)\\ &\quad= \sum_{k=0}^{M-1}\frac{e^{-\frac{2\pi itk}{M}}}{M}\sum_{b=1}^{N^2}\frac{e^{\frac{2\pi ibd}{N^2}}}{N^2}e^{\frac{2\pi inmk}{N}}\,\mathrm{Tr}_{\mathcal{H}_{g^m}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,m}(g)^{-b}\tilde{\rho}_{g,m}(h)^k(-1)^{F+\bar{F}}\big)\\ &\quad= \sum_{k=0}^{M-1}\frac{e^{-\frac{2\pi i(t-sm)k}{M}}}{M}\sum_{b=1}^{N^2}\frac{e^{\frac{2\pi ibd}{N^2}}}{N^2}\,\mathrm{Tr}_{\mathcal{H}_{g^m}(\frac{md}{N^2},\ell)}\big(\tilde{\rho}_{g,m}(g)^{-b}\tilde{\rho}_{g,m}(h)^k(-1)^{F+\bar{F}}\big)\\ &\quad= \hat{c}_{g,h}(m,d,\ell,t-sm)\ .\end{aligned} \tag{D.60}$$

Using this identity, we obtain

$$\begin{aligned}
\Phi_{g,h}(\sigma,\tau,z) &= pq^{\frac{1}{N^2}}y\prod_{t=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N^2}}y^{\ell}p^{d})^{\hat{c}_{g,h}(d,m,\ell,t)} \\
&= pq^{\frac{1}{N^2}}y\prod_{t=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi it}{M}}q^{\frac{m}{N^2}}y^{\ell}p^{d})^{\hat{c}_{g,h}(m,d,\ell,t-sm)} \\
&= pq^{\frac{1}{N^2}}y\prod_{t'=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi ism}{M}}e^{\frac{2\pi it'}{M}}q^{\frac{m}{N^2}}y^{\ell}p^{d})^{\hat{c}_{g,h}(m,d,\ell,t')} \\
&= pq^{\frac{1}{N^2}}y\prod_{t'=1}^{M}\prod_{(d,m,\ell)>0}(1-e^{\frac{2\pi inm}{N}}e^{\frac{2\pi it'}{M}}q^{\frac{m}{N^2}}y^{\ell}p^{d})^{\hat{c}_{g,h}(m,d,\ell,t')} \\
&= e^{-\frac{2\pi in}{N}}\Phi_{g,h}(\frac{\tau}{N^2}+\frac{n}{N},N^2\sigma,z) \\
&= e^{\frac{2\pi in}{N}}\Phi_{g,h}(\frac{\tau}{N^2},N^2\sigma-Nn,z)\ ,
\end{aligned} \tag{D.61}$$

where we used the property that the exponent $\hat{c}_{g,h}(m,d,\ell,t')$ is non-zero only for $d\equiv -m \mod N^2$, so that

$$e^{-\frac{2\pi ix}{N^2}}\Phi_{g,h}(\sigma+\frac{x}{N^2},\tau,z)=e^{\frac{2\pi ix}{N^2}}\Phi_{g,h}(\sigma,\tau-x,z)\ ,\qquad x\in\mathbb{Z}\ . \tag{D.62}$$

The identity (D.61) can be simplified by a suitable choice of the 3-cocycle. Let us shift the cocycle α by a 3-coboundary $\partial\beta$ such that $\frac{\beta(g,h)}{\beta(h,g)}=e^{\frac{2\pi ir}{N^2}}$ for some $r\in\mathbb{Z}$, so that, by (3.56), the new form Φ' satisfies

$$\begin{aligned}
\Phi'_{g,h}(\sigma,\tau,z) &= e^{-\frac{2\pi ir}{N^2}}\Phi_{g,h}(\sigma+\frac{r}{N^2},\tau,z)=e^{\frac{2\pi i(Nn-r)}{N^2}}\Phi_{g,h}(\frac{\tau}{N^2},N^2\sigma+r-Nn,z) \\
&= e^{\frac{2\pi i(r-Nn)}{N^2}}\Phi_{g,h}(\frac{\tau}{N^2}+\frac{Nn-r}{N^2},N^2\sigma,z)=e^{\frac{2\pi i(2r-Nn)}{N^2}}\Phi'_{g,h}(\frac{\tau}{N^2}+\frac{Nn-2r}{N^2},N^2\sigma,z)\ .
\end{aligned} \tag{D.63}$$

By choosing $r=\frac{n(N+N^2)}{2}$, we finally obtain

$$\Phi'_{g,h}(\sigma,\tau,z)=\Phi'_{g,h}(\frac{\tau}{N^2}-n,N^2\sigma,z)=\Phi'_{g,h}(\frac{\tau}{N^2},N^2\sigma,z)\ . \tag{D.64}$$

E Modular forms and Jacobi forms

The Dedekind η function and Jacobi theta functions are defined as

$$\begin{aligned}
\eta(\tau) &= q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n) \\
\vartheta_1(\tau,z) &= -iq^{\frac{1}{8}}y^{\frac{1}{2}} \prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^{n-1}) \\
\vartheta_2(\tau,z) &= 2\,q^{\frac{1}{8}}\cos(\pi z) \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^n)(1+y^{-1}q^n) \\
\vartheta_3(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^{n-1/2})(1+y^{-1}q^{n-1/2}) \\
\vartheta_4(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1-yq^{n-1/2})(1-y^{-1}q^{n-1/2}) \quad .
\end{aligned} \tag{E.1}$$

The standard weak Jacobi forms $\phi_{0,1}$ and $\phi_{-2,1}$ of index 1 and weight 0 and 2 can be defined as [57]

$$\phi_{0,1}(\tau,z) = 4\sum_{i=2}^{4}\frac{\vartheta_i(\tau,z)^2}{\vartheta_i(\tau,0)^2}\ , \qquad \phi_{-2,1}(\tau,z) = -\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^6}\ . \tag{E.2}$$

Every weak Jacobi forms of weight 0 and index 1 under a congruence subgroup $\Gamma \subseteq SL(2,\mathbb{Z})$ is given by

$$A\phi_{0,1}(\tau,z) + F(\tau)\phi_{-2,1}(\tau,z)\ , \tag{E.3}$$

where A is a constant and $F(\tau)$ is a modular form of weight 2 for Γ.

The Eisenstein series

$$\psi^{(N)} = q\frac{\partial}{\partial q}\log\frac{\eta(N\tau)}{\eta(\tau)} = E_2(\tau) - NE_2(N\tau)\ , \tag{E.4}$$

where

$$E_2(\tau) = -\frac{1}{24} + \sum_{n=1}^{\infty}\Big(\sum_{d|n} d\Big)q^n\ , \tag{E.5}$$

are modular forms of weight 2 under $\Gamma_0(N) \subset SL(2,\mathbb{Z})$. We will also need the newforms

$$f_{23,a}(\tau) = q - q^3 - q^4 - 2q^6 + 2q^7 - q^8 + 2q^9 + 2q^{10} + \cdots \tag{E.6}$$

$$f_{23,b}(\tau) = -q^2 + 2q^3 + q^4 - 2q^5 - q^6 - 2q^7 + 2q^8 + 2q^{10} + \cdots \tag{E.7}$$

that are modular (cusp) forms of weight 2 for $\Gamma_0(23)$.

F Tables

In the following tables, we collect information about the 55 abelian subgroups $\langle g, h\rangle$ of M_{24} and the corresponding twisted twining genera.

Table 1 describes the 21 cyclic subgroups $\langle g\rangle$, $g \in M_{24}$ and the corresponding twining genera. For each group we report the conjugacy class of the generators, the orbits in the 24-dim representation, the twisted twining genera for some pairs of elements in the group. The twining genus $\phi_{e,g}$ is a weak Jacobi form of weight 0 and index 1 for the group

$$\Gamma_0(N) := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid c \equiv 0 \mod N\} , \tag{F.1}$$

where N is the order of the g. Thus, $\phi_{e,g}$ is given by $\frac{\mathrm{Tr}_{\mathbf{24}}(g)}{12}\phi_{0,1}(\tau,z) + F_{e,g}(\tau)\phi_{-2,1}(\tau,z)$, where $\phi_{0,1}(\tau,z)$, $\phi_{-2,1}(\tau,z)$ are the standard generators of the ring of weak Jacobi forms of index 1 and $\mathrm{Tr}_{\mathbf{24}}(g)$ and $F_{e,g}(\tau)$ are reported in the last two columns. The table is divided into two parts, the first containing the groups whose generators are in some M_{23} subgroup of M_{24} (equivalently, the groups have some fixed points in the 24-dimensional permutation representation).

Table 2 contains information about the 34 conjugacy classes of non-cyclic abelian groups of M_{24} (see [17]). For each such $\langle g, h\rangle \subset M_{24}$ we have described the structure as an abelian group, i.e. as $\mathbb{Z}_m \times \mathbb{Z}_n$, the M_{24} classes of all its elements (excluding the identity), the order of the centralizer $C(g,h)$ and its index $|N(g,h)|/|C(g,h)|$ in the normalizer of $\langle g, h\rangle$ in M_{24}, and the lengths of the orbits of $\langle g, h\rangle \subset M_{24}$ when acting as a group of permutations of 24 objects. Furthermore, we give a classification of all modular groups $\Gamma_{g,h} \subset PSL(2,\mathbb{Z})$ and twisted twining genera $\phi_{g,h}$ for commuting pairs $g, h \in M_{24}$. The latter functions are denoted as $\phi_{[g],[h]}$, where the first subscript is the M_{24}-class of g and the second is $C_{M_{24}}(g)$-class of h; the names of the classes follow the conventions of [17]. Finally, for each of the 34 groups we provide the explicit expressions of the twisted twining genera $\phi_{g,h}$ and $\mathcal{T}_L^{\alpha}\phi_{g,h}$, $L = 2, 3, 4$, for a pair (g,h) of generators and for a choice of cocycle α satisfying (D.17) and (D.21) (when g is in $M_{23} \subset M_{24}$) or (D.27) (otherwise).

Group 27 is the only case where the pairs (g,h) and (g^{-1},h^{-1}) are not conjugated within M_{24}. Thus, in this case, charge conjugation gives the identities $\phi_{\mathrm{2B},\mathrm{8A}_1} = \phi_{\mathrm{2B},\mathrm{8A}_2}$, $\phi_{\mathrm{8A},\mathrm{2B}_1} = \phi_{\mathrm{8A},\mathrm{2B}_2}$, and so on, and the respective functions are denoted in the following tables by $\phi_{\mathrm{2B},\mathrm{8A}_{1,2}}$, $\phi_{\mathrm{8A},\mathrm{2B}_{1,2}}$, etcetera.

Most of the modular groups $\Gamma_{g,h}$ are of the form $\Gamma(1) = SL(2,\mathbb{Z})$, $\Gamma_0(N)$ or conjugates of $\Gamma_0(N)$ in $SL(2,\mathbb{Z})$. The exceptions are the group in case 32, where

$$\Gamma_{\mathrm{2B},\mathrm{10A}} = \bigcup_{i\in\mathbb{Z}/3\mathbb{Z}, j\in\mathbb{Z}/4\mathbb{Z}} \left(\begin{smallmatrix} 1 & 1 \\ -5 & -4 \end{smallmatrix}\right)^i \left(\begin{smallmatrix} -3 & -1 \\ 10 & 3 \end{smallmatrix}\right)^j \Gamma_{2,10} , \tag{F.2}$$

is a subgroup of index 12 in $SL(2,\mathbb{Z})$ and

$$\Gamma_{2,10} := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid a \equiv 1,\ b \equiv 0 \mod 2,\ c \equiv 0,\ d \equiv 1 \mod 10\}\ , \tag{F.3}$$

is the group of elements $\gamma \in SL(2,\mathbb{Z})$ such that $(g,h)\cdot\gamma = (g,h)$; the group in case 12, with

$$\Gamma_{\mathrm{2A,4A}} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid b \equiv 0 \mod 2, c \equiv 0 \mod 4\}\ , \tag{F.4}$$

which is a conjugate of $\Gamma_0(8)$ in $SL(2,\mathbb{R})$; and the group in case 22, with

$$\Gamma_{\mathrm{4A,4C}} = \langle \left(\begin{smallmatrix} -1 & 1 \\ -2 & 1 \end{smallmatrix}\right)\ ,\ \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)\ ,\ \left(\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right)\ ,\ \left(\begin{smallmatrix} 3 & -2 \\ -4 & 3 \end{smallmatrix}\right) \rangle\ . \tag{F.5}$$

g	Orbits on **24**	Functions	$\mathrm{Tr}_{\mathbf{24}}(g)$	$F_{e,g}(\tau)$
1A	1^{24}		24	0
2A	$1^8\cdot 2^8$	$\phi_{2A,2A_1}$	8	$32\,\psi^{(2)}(\tau)$
3A	$1^6\cdot 3^6$	$\phi_{3A,3A_{1,2}}$	6	$18\,\psi^{(3)}(\tau)$
4B	$1^4\cdot 2^2\cdot 4^4$	$\phi_{4B,4B_{2,6}},\ \phi_{4B,2A_1},\ \phi_{2A,4B_1}$	4	$-8\,\psi^{(2)}(\tau)+16\,\psi^{(4)}(\tau)$
5A	$1^4\cdot 5^4$	$\phi_{5A,5A_{1,2,3,4}}$	4	$10\,\psi^{(5)}(\tau)$
6A	$1^2\cdot 2^2\cdot 3^2\cdot 6^2$	$\phi_{6A,6A_{5,6}},\ \phi_{6A,3A_{1,2}},\ \phi_{6A,2A_3},$ $\phi_{3A,6A_{1,2}},\phi_{3A,2A_1},\ \phi_{2A,6A_1},\ \phi_{2A,3A_1}$	2	$-4\,\psi^{(2)}(\tau)-6\,\psi^{(3)}(\tau)+12\,\psi^{(6)}(\tau)$
7AB	$1^3\cdot 7^3$	$\phi_{7A,7A_{1,2,3}},\ \phi_{7A,7B_{1,2,3}}$	3	$7\,\psi^{(7)}(\tau)$
8A	$1^2\cdot 2\cdot 4\cdot 8^2$	$\phi_{8A,8A_{1,3,4,5}},\ \phi_{8A,4B_{2,4}},$ $\phi_{8A,2A_1},\ \phi_{4B,8A_{1,4}},\ \phi_{2A,8A_1}$	2	$-4\,\psi^{(4)}(\tau)+8\,\psi^{(8)}(\tau)$
11A	$1^2\cdot 11^2$		2	$\frac{22}{5}\big(\psi^{(11)}(\tau)-\eta(\tau)^2\eta(11\tau)^2\big)$
14AB	$1\cdot 2\cdot 7\cdot 14$	$\phi_{2A,7AB_1},\ \phi_{2A,14AB_1},\ \phi_{7AB,2A_1},$ $\phi_{7AB,14A_{1,2,3}},\ \phi_{7AB,14B_{1,2,3}}$	1	$\frac{1}{3}\big(-2\,\psi^{(2)}(\tau)-7\,\psi^{(7)}(\tau)+14\,\psi^{(14)}(\tau)$ $-14\,\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau)\big)$
15AB	$1\cdot 3\cdot 5\cdot 15$	$\phi_{3A,5A_{1,2}},\ \phi_{3A,15A_{1,2}},\ \phi_{3A,15B_{1,2}}$ $\phi_{5A,3A_{1,2}},\ \phi_{5A,15A_{1,2,3,4}},\ \phi_{5A,15B_{1,2,3,4}}$	1	$\frac{1}{4}\big(-3\,\psi^{(3)}(\tau)-5\,\psi^{(5)}(\tau)+15\,\psi^{(15)}(\tau)$ $-15\,\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau)\big)$
23AB	$1\cdot 23$		1	$\frac{23}{11}\big(\psi^{(23)}(\tau)-f_{23,a}(\tau)+3f_{23,b}(\tau)\big)$
2B	2^{12}	$\phi_{2B,2B_3}$	0	$2\,\frac{\eta(\tau)^8}{\eta(2\tau)^4}$
3B	3^8	$\phi_{3B,3B_{1,2}}$	0	$2\,\frac{\eta(\tau)^6}{\eta(3\tau)^2}$
4A	$2^4\cdot 4^4$	$\phi_{4A,4A_{6,8}},\ \phi_{4A,2A_4},\ \phi_{2A,4A_1}$	0	$2\,\frac{\eta(2\tau)^8}{\eta(4\tau)^4}$
4C	4^6	$\phi_{4C,4C_{1,2}},\ \phi_{4C,2B_1},\ \phi_{2B,4C_1}$	0	$2\,\frac{\eta(\tau)^4\,\eta(2\tau)^2}{\eta(4\tau)^2}$
6B	6^4	$\phi_{6B,6B_{4,6}},\ \phi_{6B,3B_{1,2}},\ \phi_{6B,2B_3},$ $\phi_{3B,6B_{1,2}},\phi_{3B,2B_1},\ \phi_{2B,6B_1},\ \phi_{2B,3B_1}$	0	$2\,\frac{\eta(\tau)^2\,\eta(2\tau)^2\,\eta(3\tau)^2}{\eta(6\tau)^2}$
10A	$2^2\cdot 10^2$	$\phi_{10A,10A_{4,6,8,12}},\ \phi_{10A,2B_1},\ \phi_{10A,5A_{1,2,3,4}},$ $\phi_{2B,10A_2},\ \phi_{2B,5A_1},\ \phi_{5A,10A_{1,2,3,4}},\ \phi_{5A,2B_1}$	0	$2\,\frac{\eta(\tau)^3\,\eta(2\tau)\,\eta(5\tau)}{\eta(10\tau)}$
12A	$2\cdot 4\cdot 6\cdot 12$	$\phi_{2A,12A_1},\ \phi_{3A,12A_{1,2}},\ \phi_{3A,4A_1},\ \phi_{4A,12A_{1,2}},$ $\phi_{4A,6A_1},\ \phi_{4A,3A_1},\ \phi_{6A,12A_{1,2}},\ \phi_{6A,4A_1}$	0	$2\,\frac{\eta(\tau)^3\,\eta(4\tau)^2\,\eta(6\tau)^3}{\eta(2\tau)\,\eta(3\tau)\,\eta(12\tau)^2}$
12B	12^2	$\phi_{2B,12B_1},\ \phi_{3B,12B_{1,2}},\ \phi_{3B,4C_1},\ \phi_{4C,12B_{1,2}},$ $\phi_{4C,6B_1},\ \phi_{4C,3B_1},\ \phi_{6B,12B_{1,2}},\ \phi_{6B,4C_1}$	0	$2\,\frac{\eta(\tau)^4\,\eta(4\tau)\,\eta(6\tau)}{\eta(2\tau)\,\eta(12\tau)}$
21AB	$3\cdot 21$	$\phi_{3B,21A_{1,2}},\ \phi_{3B,21B_{1,2}},\ \phi_{3B,7A_1},\ \phi_{3B,7B_1},$ $\phi_{7AB,21A_{1,2,3}},\ \phi_{7AB,21B_{1,2,3}},\ \phi_{7AB,3B_1}$	0	$\frac{7}{3}\,\frac{\eta(\tau)^3\,\eta(7\,\tau)^3}{\eta(3\,\tau)\,\eta(21\,\tau)}-\frac{1}{3}\,\frac{\eta(\tau)^6}{\eta(3\,\tau)^2}$

Table 1: The cyclic subgroups of M_{24}.

#	Structure	Elements	$\lvert C(g,h)\rvert$	$\frac{\lvert N(g,h)\rvert}{\lvert C(g,h)\rvert}$	Orbits on **24**	Max subgr.	Functions	$\Gamma_{g,h}$
1.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	1536	6	$1^8 \cdot 4^4$		$\phi_{2\mathrm{A},2\mathrm{A}_2}$	$\Gamma(1)$
2.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	1536	6	2^{12}		$\phi_{2\mathrm{A},2\mathrm{A}_3}$	$\Gamma(1)$
3.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	128	6	$1^4 \cdot 2^6 \cdot 4^2$		$\phi_{2\mathrm{A},2\mathrm{A}_5}$	$\Gamma(1)$
4.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	3840	6	4^6		$\phi_{2\mathrm{B},2\mathrm{B}_2}$	$\Gamma(1)$
5.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	96	6	4^6		$\phi_{2\mathrm{B},2\mathrm{B}_4}$	$\Gamma(1)$
6.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	64	6	4^6		$\phi_{2\mathrm{B},2\mathrm{B}_6}$	$\Gamma(1)$
7.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^2(2\mathrm{B})$	256	2	$2^8 \cdot 4^2$		$\phi_{2\mathrm{B},2\mathrm{A}_1},\ \phi_{2\mathrm{A},2\mathrm{B}_3},\ \phi_{2\mathrm{A},2\mathrm{A}_4}$	$\Gamma_0(2)$
8.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})(2\mathrm{B})^2$	512	2	$2^4 \cdot 4^4$		$\phi_{2\mathrm{A},2\mathrm{B}_1},\ \phi_{2\mathrm{B},2\mathrm{A}_2},\ \phi_{2\mathrm{B},2\mathrm{B}_1}$	$\Gamma_0(2)$
9.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})(2\mathrm{B})^2$	128	2	$2^4 \cdot 4^4$		$\phi_{2\mathrm{A},2\mathrm{B}_2},\ \phi_{2\mathrm{B},2\mathrm{A}_3},\ \phi_{2\mathrm{B},2\mathrm{B}_5}$	$\Gamma_0(2)$
10.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^4$	64	8	$2^4 \cdot 8^2$	1	$\phi_{2\mathrm{A},4\mathrm{A}_2},\ \phi_{4\mathrm{A},2\mathrm{A}_2},\ \phi_{4\mathrm{A},4\mathrm{A}_3}$	$\Gamma_0(2)$
11.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^4$	64	8	4^6	2	$\phi_{2\mathrm{A},4\mathrm{A}_3},\ \phi_{4\mathrm{A},2\mathrm{A}_3},\ \phi_{4\mathrm{A},4\mathrm{A}_7}$	$\Gamma_0(2)$
12.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^2(2\mathrm{B})$ $(4\mathrm{A})^2(4\mathrm{B})^2$	32	2	$2^2 \cdot 4^3 \cdot 8^1$	7	$\phi_{2\mathrm{A},4\mathrm{A}_4}, \phi_{2\mathrm{A},4\mathrm{B}_4}, \phi_{2\mathrm{B},4\mathrm{A}_1},$ $\phi_{2\mathrm{B},4\mathrm{B}_1}, \phi_{4\mathrm{A},2\mathrm{A}_1}, \phi_{4\mathrm{A},2\mathrm{B}_1},$ $\phi_{4\mathrm{B},2\mathrm{A}_4}, \phi_{4\mathrm{B},2\mathrm{B}_2}, \phi_{4\mathrm{B},4\mathrm{A}_1},$ $\phi_{4\mathrm{B},4\mathrm{A}_2}, \phi_{4\mathrm{A},4\mathrm{B}_3}, \phi_{4\mathrm{A},4\mathrm{B}_4}$	$\Gamma_{2\mathrm{A},4\mathrm{A}}$
13.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{A})^4$	64	8	$4^2 \cdot 8^2$	8	$\phi_{2\mathrm{B},4\mathrm{A}_2},\ \phi_{4\mathrm{A},2\mathrm{B}_3},\ \phi_{4\mathrm{A},4\mathrm{A}_5}$	$\Gamma_0(2)$
14.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{A})^4$	32	8	$4^2 \cdot 8^2$	8	$\phi_{2\mathrm{B},4\mathrm{A}_3},\ \phi_{4\mathrm{A},2\mathrm{B}_2},\ \phi_{4\mathrm{A},4\mathrm{A}_2}$	$\Gamma_0(2)$
15.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{C})^4$	32	4	$4^2 \cdot 8^2$	8	$\phi_{2\mathrm{A},4\mathrm{C}_1}, \phi_{4\mathrm{C},2\mathrm{A}_2}, \phi_{4\mathrm{C},2\mathrm{B}_3},$ $\phi_{2\mathrm{B},4\mathrm{C}_2}, \phi_{4\mathrm{C},4\mathrm{C}_4}, \phi_{4\mathrm{C},4\mathrm{C}_6}$	$\Gamma_0(4)$
16.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{C})^4$	16	4	$4^2 \cdot 8^2$	9	$\phi_{2\mathrm{A},4\mathrm{C}_2}, \phi_{4\mathrm{C},2\mathrm{A}_1}, \phi_{4\mathrm{C},2\mathrm{B}_2},$ $\phi_{2\mathrm{B},4\mathrm{C}_3}, \phi_{4\mathrm{C},4\mathrm{C}_3}, \phi_{4\mathrm{C},4\mathrm{C}_5}$	$\Gamma_0(4)$
17.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	64	8	$1^4 \cdot 2^2 \cdot 8^2$	1	$\phi_{2\mathrm{A},4\mathrm{B}_2},\ \phi_{4\mathrm{B},2\mathrm{A}_2},\ \phi_{4\mathrm{B},4\mathrm{B}_8}$	$\Gamma_0(2)$
18.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	64	8	$2^4 \cdot 4^4$	2	$\phi_{2\mathrm{A},4\mathrm{B}_3},\ \phi_{4\mathrm{B},2\mathrm{A}_5},\ \phi_{4\mathrm{B},4\mathrm{B}_9}$	$\Gamma_0(2)$
19.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	16	8	$1^2 \cdot 2^3 \cdot 4^2 \cdot 8^1$	3	$\phi_{2\mathrm{A},4\mathrm{B}_5},\ \phi_{4\mathrm{B},2\mathrm{A}_3},\ \phi_{4\mathrm{B},4\mathrm{B}_3}$	$\Gamma_0(2)$

Table 2: The non-cyclic abelian groups $\langle g, h\rangle$.

#	Structure	Elements	$\lvert C(g,h)\rvert$	$\frac{\lvert N(g,h)\rvert}{\lvert C(g,h)\rvert}$	Orbits on **24**	Max subgr.	Functions	$\Gamma_{g,h}$
20.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2A)(2B)^2(4B)^4$	64	8	$2^4 \cdot 8^2$	8	$\phi_{\mathrm{2B,4B}_2}$, $\phi_{\mathrm{4B,2B}_3}$, $\phi_{\mathrm{4B,4B}_1}$	$\Gamma_0(2)$
21.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2A)(2B)^2(4B)^4$	16	8	$2^4 \cdot 8^2$	9	$\phi_{\mathrm{2B,4B}_3}$, $\phi_{\mathrm{4B,2B}_1}$, $\phi_{\mathrm{4B,4B}_5}$	$\Gamma_0(2)$
22.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2A)(2B)^2$ $(4A)^4(4C)^8$	16	16	$8^1 \cdot 16^1$	13, 15, 15	$\phi_{\mathrm{4A,4C}_1}$, $\phi_{\mathrm{4A,4C}_2}$, $\phi_{\mathrm{4C,4A}_1}$, $\phi_{\mathrm{4C,4A}_2}$, $\phi_{\mathrm{4C,4C}_7}$, $\phi_{\mathrm{4C,4C}_8}$	$\Gamma_{\mathrm{4A,4C}}$
23.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2A)^3(4A)^8(4B)^4$	16	32	$2^2 \cdot 4^1 \cdot 16^1$	10, 10, 17	$\phi_{\mathrm{4B,4A}_3}$, $\phi_{\mathrm{4A,4B}_1}$, $\phi_{\mathrm{4A,4A}_1}$	$\Gamma_0(2)$
24.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2A)^3(4A)^8(4B)^4$	16	32	$4^2 \cdot 8^2$	11, 11, 18	$\phi_{\mathrm{4B,4A}_4}$, $\phi_{\mathrm{4A,4B}_2}$, $\phi_{\mathrm{4A,4A}_4}$	$\Gamma_0(2)$
25.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2A)^3(4B)^{12}$	16	96	$1^4 \cdot 4^1 \cdot 16^1$	17, 17, 17	$\phi_{\mathrm{4B,4B}_4}$	$\Gamma(1)$
26.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2A)^3(4B)^{12}$	16	96	4^6	18, 18, 18	$\phi_{\mathrm{4B,4B}_7}$	$\Gamma(1)$
27.	$\mathbb{Z}_2 \times \mathbb{Z}_8$	$(2A)(2B)^2$ $(4B)^4(8A)^8$	16	8	$2^2 \cdot 4^1 \cdot 16^1$	20	$\phi_{\mathrm{2B,8A}_{1,2}}$, $\phi_{\mathrm{8A,2B}_{1,2}}$, $\phi_{\mathrm{4B,8A}_{2,3}}$, $\phi_{\mathrm{8A,4B}_{1,3}}$, $\phi_{\mathrm{8A,8A}_{2,8}}$, $\phi_{\mathrm{8A,8A}_{6,7}}$	$\Gamma_0(4)$
28.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2A)^3(3A)^2(6A)^6$	12	12	$1^2 \cdot 3^2 \cdot 4^1 \cdot 12^1$	1	$\phi_{\mathrm{2A,6A}_2}$, $\phi_{\mathrm{6A,2A}_1}$, $\phi_{\mathrm{6A,6A}_1}$, $\phi_{\mathrm{6A,6A}_2}$	$\Gamma_0(3)$
29.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2A)^3(3A)^2(6A)^6$	12	12	$2^3 \cdot 6^3$	2	$\phi_{\mathrm{2A,6A}_3}$, $\phi_{\mathrm{6A,2A}_2}$, $\phi_{\mathrm{6A,6A}_3}$, $\phi_{\mathrm{6A,6A}_4}$	$\Gamma_0(3)$
30.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2B)^3(3B)^2(6B)^6$	12	12	12^2	4	$\phi_{\mathrm{2B,6B}_2}$, $\phi_{\mathrm{6B,2B}_1}$, $\phi_{\mathrm{6B,6B}_2}$, $\phi_{\mathrm{6B,6B}_5}$	$\Gamma_0(3)$
31.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2B)^3(3B)^2(6B)^6$	12	12	12^2	5	$\phi_{\mathrm{2B,6B}_3}$, $\phi_{\mathrm{6B,2B}_2}$, $\phi_{\mathrm{6B,6B}_1}$, $\phi_{\mathrm{6B,6B}_3}$	$\Gamma_0(3)$
32.	$\mathbb{Z}_2 \times \mathbb{Z}_{10}$	$(2B)^3(5A)^4(10A)^{12}$	20	12	$4^1 \cdot 20^1$	4	$\phi_{\mathrm{2B,10A}_1}$, $\phi_{\mathrm{2B,10A}_3}$, $\phi_{\mathrm{10A,10A}_1}$, $\phi_{\mathrm{10A,10A}_2}$, $\phi_{\mathrm{10A,10A}_3}$, $\phi_{\mathrm{10A,10A}_5}$, $\phi_{\mathrm{10A,10A}_7}$, $\phi_{\mathrm{10A,10A}_9}$, $\phi_{\mathrm{10A,10A}_{10}}$, $\phi_{\mathrm{10A,10A}_{11}}$, $\phi_{\mathrm{10A,2B}_2}$, $\phi_{\mathrm{10A,2B}_3}$	$\Gamma_{\mathrm{2B,10A}}$
33.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$(3A)^8$	9	48	$1^3 \cdot 3^4 \cdot 9^1$		$\phi_{\mathrm{3A,3A}_3}$	$\Gamma(1)$
34.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$(3A)^2(3B)^6$	9	12	$3^2 \cdot 9^2$		$\phi_{\mathrm{3A,3B}_1}$, $\phi_{\mathrm{3B,3A}_1}$, $\phi_{\mathrm{3B,3B}_3}$, $\phi_{\mathrm{3B,3B}_4}$	$\Gamma_0(3)$

Table 2 (continued)

#	(g,h)	$\phi_{g,h}$	$2\mathcal{T}_2^{\alpha}\phi_{g,h}$	$3\mathcal{T}_3^{\alpha}\phi_{g,h}$	$4\mathcal{T}_4^{\alpha}\phi_{g,h}$
1.	$(2A,2A_2)$	0	$\phi_{2A,e}(2\tau,2z)+\phi_{e,2A}\left(\frac{\tau}{2},z\right)+\phi_{e,2A}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,e}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,2A}\left(\frac{\tau+i}{4},z\right)$
2.	$(2A,2A_3)$				
3.	$(2A,2A_5)$				
4.	$(2B,2B_2)$	0	$-\phi_{2B,e}(2\tau,2z)+\phi_{e,2B}\left(\frac{\tau}{2},z\right)-\phi_{e,2B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2B,e}(4\tau,4z)-\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,e}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,2B}\left(\frac{\tau+i}{4},z\right)$
5.	$(2B,2B_4)$				
6.	$(2B,2B_6)$				
7.	$(2A,2A_4)$	0	$-\phi_{2A,e}(2\tau,2z)+\phi_{e,2A}\left(\frac{\tau}{2},z\right)+\phi_{e,2B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2A}\left(\frac{2\tau+1}{2},2z\right)+\phi_{e,e}(\tau,2z)$ $+\sum_{i=0}^{1}\left(\phi_{e,2A}\left(\frac{\tau+2i}{4},z\right)+\phi_{e,2B}\left(\frac{\tau+2i+1}{4},z\right)\right)$
8.	$(2A,2B_1)$	0	$\phi_{2A,e}(2\tau,2z)+\phi_{e,2B}\left(\frac{\tau}{2},z\right)+\phi_{e,2B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,e}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,2B}\left(\frac{\tau+i}{4},z\right)$
9.	$(2A,2B_2)$				
10.	$(2A,4A_2)$	0	$\phi_{e,4A}\left(\frac{\tau}{2},z\right)+\phi_{e,4A}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4A}\left(\frac{\tau+i}{4},z\right)$
11.	$(2A,4A_3)$				
12.	$(2A,4B_4)$	0	$\phi_{e,4A}\left(\frac{\tau+1}{2},z\right)+\phi_{e,4B}\left(\frac{\tau}{2},z\right)$	0	$-\phi_{2A,e}(4\tau,4z)+\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)+\phi_{e,2A}(\tau,2z)$ $+\sum_{i=0}^{1}\left(\phi_{e,4B}\left(\frac{\tau+2i}{4},z\right)+\phi_{e,4A}\left(\frac{\tau+2i+1}{4},z\right)\right)$
13.	$(2B,4A_2)$	$4\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	$\phi_{e,4A}\left(\frac{\tau}{2},z\right)-\phi_{e,4A}\left(\frac{\tau+1}{2},z\right)$	$\phi_{2B,4A_2}(3\tau,3z)+\phi_{2B,4A_2}\left(\frac{\tau}{3},z\right)$ $+i\phi_{2B,4A_2}\left(\frac{\tau+1}{3},z\right)-\phi_{2B,4A_2}\left(\frac{\tau+2}{3},z\right)$	$\phi_{2B,e}(4\tau,4z)-\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4A}\left(\frac{\tau+i}{4},z\right)$
14.	$(2B,4A_3)$	0	$\phi_{e,4A}\left(\frac{\tau}{2},z\right)-\phi_{e,4A}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2B,e}(4\tau,4z)-\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4A}\left(\frac{\tau+i}{4},z\right)$
15.	$(2A,4C_1)$	0	$\phi_{e,4C}\left(\frac{\tau}{2},z\right)+\phi_{e,4C}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2B}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4C}\left(\frac{\tau+i}{4},z\right)$
16.	$(2A,4C_2)$				
17.	$(2A,4B_2)$	0	$\phi_{e,4B}\left(\frac{\tau}{2},z\right)+\phi_{e,4B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,e}(4\tau,4z)+\phi_{e,2A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4B}\left(\frac{\tau+i}{4},z\right)$
18.	$(2A,4B_3)$				
19.	$(2A,4B_5)$				

Table 2 (continued)

#	(g,h)	$\phi_{g,h}$	$2\mathcal{T}_2^\alpha\phi_{g,h}$	$3\mathcal{T}_3^\alpha\phi_{g,h}$	$4\mathcal{T}_4^\alpha\phi_{g,h}$
20.	$(2B,4B_2)$	0	$\phi_{e,4B}\left(\frac{\tau}{2},z\right)-\phi_{e,4B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2B,e}(4\tau,4z)-\phi_{e,2B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,2A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,4B}\left(\frac{\tau+i}{4},z\right)$
21.	$(2B,4B_3)$				
22.	$(4A,4C_1)$	0	$\phi_{4A,2B_3}(2\tau,2z)$	0	$-\phi_{4A,e}(4\tau,4z)+\phi_{2A,4A_1}\left(\frac{2\tau+1}{2},2z\right)$ $+\sum_{i=0}^{3}(-1)^i\phi_{e,4C}\left(\frac{\tau+i}{4},z\right)$
23.	$(4B,4A_3)$	$2\sqrt{2}\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	0	$\phi_{4B,4A_3}(3\tau,3z)+\phi_{4B,4A_3}\left(\frac{\tau}{3},z\right)$ $+i\phi_{4B,4A_3}\left(\frac{\tau+1}{3},z\right)-\phi_{4B,4A_3}\left(\frac{\tau+2}{3},z\right)$	$\phi_{4B,e}(4\tau,4z)+\phi_{2A,4B_1}\left(\frac{2\tau+1}{2},2z\right)$ $+\sum_{i=0}^{3}\phi_{e,4A}\left(\frac{\tau+i}{4},z\right)$
24.	$(4B,4A_4)$				
25.	$(4B,4B_4)$	0	0	0	$\phi_{4B,e}(4\tau,4z)+\phi_{2A,4B_1}\left(\frac{2\tau+1}{2},2z\right)$ $+\sum_{i=0}^{3}\phi_{e,4B}\left(\frac{\tau+i}{4},z\right)$
26.	$(4B,4B_7)$				
27.	$(2B,8A_1)$	$2\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	$\phi_{e,8A}\left(\frac{\tau}{2},z\right)-\phi_{e,8A}\left(\frac{\tau+1}{2},z\right)$	$\phi_{2B,8A_{1,2}}(3\tau,3z)+\phi_{2B,8A_{1,2}}\left(\frac{\tau}{3},z\right)$ $+i\phi_{2B,8A_{1,2}}\left(\frac{\tau+1}{3},z\right)-\phi_{2B,8A_{1,2}}\left(\frac{\tau+2}{3},z\right)$	$-\phi_{e,4B}\left(\frac{2\tau+1}{2},2z\right)+\phi_{e,4B}\left(\tau,2z\right)$ $+\sum_{i=0}^{3}\phi_{e,8A}\left(\frac{\tau+i}{4},z\right)$
28.	$(2A,6A_2)$	0	$\phi_{2A,3A_1}(2\tau,2z)+\phi_{e,6A}\left(\frac{\tau}{2},z\right)+\phi_{e,6A}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2A,3A}(4\tau,4z)+\phi_{e,6A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,3A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,6A}\left(\frac{\tau+i}{4},z\right)$
29.	$(2A,6A_3)$				
30.	$(2B,6B_2)$	0	$\phi_{2B,3B_1}(2\tau,2z)+\phi_{e,6B}\left(\frac{\tau}{2},z\right)-\phi_{e,6B}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2B,3B}(4\tau,4z)-\phi_{e,6B}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,3B}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,6B}\left(\frac{\tau+i}{4},z\right)$
31.	$(2B,6B_3)$				
32.	$(2B,10A_1)$	0	$\phi_{2B,5A_1}(2\tau,2z)+\phi_{e,10A}\left(\frac{\tau}{2},z\right)-\phi_{e,10A}\left(\frac{\tau+1}{2},z\right)$	0	$\phi_{2B,5A}(4\tau,4z)-\phi_{e,10A}\left(\frac{2\tau+1}{2},2z\right)$ $+\phi_{e,5A}(\tau,2z)+\sum_{i=0}^{3}\phi_{e,10A}\left(\frac{\tau+i}{4},z\right)$
33.	$(3A,3A_3)$	0	0	$\phi_{3A,e}(3\tau,3z)+\sum_{k=0}^{2}\phi_{e,3A}\left(\frac{\tau+k}{3},z\right)$	0
34.	$(3A,3B_1)$	0	0	$\phi_{3A,e}(3\tau,3z)+\sum_{k=0}^{2}\phi_{e,3B}\left(\frac{\tau+k}{3},z\right)$	0

Table 2 (continued)

Group	(g,h)	$\eta_{g,h}$	w
$1,\ 2,\ 3$	$(2A,2A_{2,3,5})$	2^{12}	6
$4,\ 5,\ 6$	$(2B,2B_{2,4,6})$	4^6	3
7	$(2A,2A_4)$	$1^4\cdot 2^2\cdot 4^4$	5
	$(2A,2B_3)$	$2^{14}/1^4$	
	$(2B,2A_1)$	$4^{14}/8^4$	
$8,\ 9$	$(2A,2B_{1,2}),\ (2B,2A_{2,3})$	$2^4 4^4$	4
	$(2B,2B_{1,5})$	$4^{16}/2^4 8^4$	
$10,\ 11$	$(2A,4A_{2,3}),\ (4A,2A_{2,3})$	4^6	3
	$(4A,4A_{3,7})$	$8^{18}/4^6\cdot 16^6$	
12	$(2A,4B_4)$	$1^2\cdot 2\cdot 4\cdot 8^2$	3
	$(2A,4A_4)$	$2^7\cdot 8^2/1^2\cdot 4$	
	$(2B,4B_1)$	$2^2\cdot 8^7/4\cdot 16^2$	
	$(2B,4A_1)$	$4^5\cdot 8^5/2^2\cdot 16^2$	
	$(4B,4A_2)$	irrational	
$13,\ 14$	$(2B,4A_{2,3}),\ (4A,2B_{3,2})$	$4^2\cdot 8^2$	2
	$(4A,4A_{5,2})$	$8^8/4^2\cdot 16^2$	
$15,\ 16$	$(2A,4C_{1,2}),\ (4C,2A_{2,1})$	$4^2\cdot 8^2$	2
	$(2B,4C_{2,3}),\ (4C,2B_{3,2})$	$8^8/4^2\cdot 16^2$	
	$(4C,4C_{4,6,3,5})$	irrational	
$17,\ 18,\ 19$	$(2A,4B_{2,3,5}),\ (4B,2A_{2,5,3})$	$2^4\cdot 4^4$	4
	$(4B,4B_{8,9,3})$	$4^{16}/2^4\cdot 8^4$	
$20,\ 21$	$(2B,4B_{2,3}),\ (4B,2B_{3,1}),\ (4B,4B_{1,5}),$	4^6	3
22	$(4A,4C_{1,2}),\ (4C,4A_{1,2})$	$8\cdot 16$	1
	$(4C,4C_{7,8})$	$16^4/8\cdot 32$	
$23,\ 24$	$(4B,4A_{3,4}),\ (4A,4B_{1,2})$	$4^2\cdot 8^2$	2
	$(4A,4A_{1,4})$	$8^8/4^2\cdot 16^2$	
$25,\ 26$	$(4B,4B_{4,7})$	4^6	3
27	$(2B,8A_{1,2}),\ (8A,2B_{1,2})$	$4^2\cdot 8^2$	2
	$(4B,8A_{2,3}),\ (8A,4B_{1,3})$	$4^2\cdot 8^2$	
	$(8A,8A_{2,6,7,8})$	$8^8/4^2\cdot 16^2$	
$28,\ 29$	$(2A,6A_{2,3}),\ (6A,2A_{1,2})$	$2^3\cdot 6^3$	3
	$(6A,6A_{1,2,3,4})$	irrational	
$30,\ 31$	$(2B,6B_{2,3}),\ (6B,2B_{1,2})$	12^2	1
	$(6B,6B_{2,5,1,3})$	irrational	
32	$(2B,10A_{1,3}),\ (10A,2B_{2,3})$	$4\cdot 20$	1
	$(10A,10A)$	irrational	
33	$(3A,3A_3)$	3^8	4
34	$(3A,3B_1),\ (3B,3A_1)$	$3^2\cdot 9^2$	2
	$(3B,3B_{3,4})$	irrational	

Table 3: The eta products of [33]. For each pair g,h, a cycle shape of the form $\prod_\ell \ell^{i(\ell)}$ is given: the corresponding eta-product is $\eta_{g,h}(\tau)=\prod_\ell \eta(\ell\tau/N\lambda)^{i(\ell)}$, where N is the order of g and λ is the length of the shortest cycle. In the last column, w is the modular weight.

$\langle g,h\rangle$	h	$\langle g,h'\rangle$	h'
$\mathbb{Z}_2A$	$1A$	$\mathbb{Z}_2A$	$1A$
$\mathbb{Z}_2A$	Q	$\mathbb{Z}_2A$	g
$\mathbb{Z}_2A$	g	$\mathbb{Z}_2A$	Q
$\mathbb{Z}_2A$	Qg	$\mathbb{Z}_2A$	Qg
$\mathbb{Z}_4A$	$4A_1$	$\mathbb{Z}_4A$	$4A_1$
$\mathbb{Z}_4A$	$4A_1$	$\mathbb{Z}_4A$	$4A_1$
$\mathbb{Z}_4B$	$4B_1$	7	$2A_4$
$\mathbb{Z}_4B$	$4B_1$	7	$2B_3$
$\mathbb{Z}_6A$	$3A_1$	$\mathbb{Z}_6A$	$3A_1$
$\mathbb{Z}_6A$	$3A_1$	$\mathbb{Z}_6A$	$6A_1$
$\mathbb{Z}_6A$	$6A_1$	$\mathbb{Z}_6A$	$3A_1$
$\mathbb{Z}_6A$	$6A_1$	$\mathbb{Z}_6A$	$6A_1$
$\mathbb{Z}_8A$	$8A_1$	12	$4B_4$
$\mathbb{Z}_8A$	$8A_1$	12	$4A_4$
$\mathbb{Z}_{12}A$	$12A_1$	$\mathbb{Z}_{12}A$	$12A_1$
$\mathbb{Z}_{12}A$	$12A_1$	$\mathbb{Z}_{12}A$	$12A_1$
$\mathbb{Z}_{14}AB$	$7A_1$	$\mathbb{Z}_{14}AB$	$7B_1$
$\mathbb{Z}_{14}AB$	$7A_1$	$\mathbb{Z}_{14}AB$	$14A_1$
$\mathbb{Z}_{14}AB$	$7B_1$	$\mathbb{Z}_{14}AB$	$7A_1$
$\mathbb{Z}_{14}AB$	$7B_1$	$\mathbb{Z}_{14}AB$	$14B_1$
$\mathbb{Z}_{14}AB$	$14A_1$	$\mathbb{Z}_{14}AB$	$7A_1$
$\mathbb{Z}_{14}AB$	$14B_1$	$\mathbb{Z}_{14}AB$	$7B_1$
$\mathbb{Z}_{14}AB$	$14A_1$	$\mathbb{Z}_{14}AB$	$14B_1$
$\mathbb{Z}_{14}AB$	$14B_1$	$\mathbb{Z}_{14}AB$	$14A_1$
1	$2A_2$	2	$2A_3$
2	$2A_3$	1	$2A_2$
3	$2A_5$	3	$2A_5$
7	$2A_4$	$\mathbb{Z}_4B$	$4B_1$
7	$2B_3$	$\mathbb{Z}_4B_1$	$4B$
8	$2B_1$	8	$2B_1$
9	$2B_2$	9	$2B_2$
10	$4A_2$	11	$4A_3$
11	$4A_3$	10	$4A_2$
12	$4A_4$	$\mathbb{Z}_8A$	$8A_1$
12	$4B_4$	$\mathbb{Z}_8A$	$8A_1$
15	$4C_1$	15	$4C_1$
16	$4C_2$	16	$4C_2$
17	$4B_2$	18	$4B_3$
18	$4B_3$	17	$4B_2$
19	$4B_5$	19	$4B_5$
28	$6A_2$	29	$6A_3$
28	$6A_2$	29	$6A_3$
29	$6A_3$	28	$6A_2$
29	$6A_3$	28	$6A_2$

Table 4: Relabeling for g in class 2A. Each line corresponds to a conjugacy class in the central extension $C^{\alpha}_{M_{24}}(g)$. For each such class, we report the corresponding class in $C_{M_{24}}(g)$ with representative h, the group $\langle g,h\rangle$ generated by g and h, the 'relabeled' class in $C_{M_{24}}(g)$ with representative h' and the group $\langle g,h'\rangle$ generated by g and h'.

$\langle g,h\rangle$	h	$\langle g,h'\rangle$	h'
$\mathbb{Z}_4B$	$1A$	$\mathbb{Z}_4B$	$1A$
$\mathbb{Z}_4B$	Q^2	$\mathbb{Z}_4B$	g^2
$\mathbb{Z}_4B$	Q	$\mathbb{Z}_4B$	g
$\mathbb{Z}_4B$	Q^3	$\mathbb{Z}_4B$	g^3
$\mathbb{Z}_4B$	g^2	$\mathbb{Z}_4B$	Q^2
$\mathbb{Z}_4B$	g^2Q^2	$\mathbb{Z}_4B$	g^2Q^2
$\mathbb{Z}_4B$	g^2Q	$\mathbb{Z}_4B$	gQ^2
$\mathbb{Z}_4B$	g^2Q^3	$\mathbb{Z}_4B$	g^3Q^2
$\mathbb{Z}_4B$	gQ^2	$\mathbb{Z}_4B$	g^2Q
$\mathbb{Z}_4B$	g	$\mathbb{Z}_4B$	Q
$\mathbb{Z}_4B$	gQ	$\mathbb{Z}_4B$	gQ
$\mathbb{Z}_4B$	gQ^3	$\mathbb{Z}_4B$	g^3Q
$\mathbb{Z}_4B$	g^3Q^2	$\mathbb{Z}_4B$	g^2Q^3
$\mathbb{Z}_4B$	g^3	$\mathbb{Z}_4B$	Q^3
$\mathbb{Z}_4B$	g^3Q	$\mathbb{Z}_4B$	gQ^3
$\mathbb{Z}_4B$	g^3Q^3	$\mathbb{Z}_4B$	g^3Q^3
$\mathbb{Z}_8A$	$8A_1$	12	$2B_2$
$\mathbb{Z}_8A$	$8A_1$	12	$2A_4$
$\mathbb{Z}_8A$	$8A_1$	12	$4A_1$
$\mathbb{Z}_8A$	$8A_1$	12	$4A_2$
$\mathbb{Z}_8A$	$8A_4$	12	$2B_2$
$\mathbb{Z}_8A$	$8A_4$	12	$2A_4$
$\mathbb{Z}_8A$	$8A_4$	12	$4A_1$
$\mathbb{Z}_8A$	$8A_4$	12	$4A_2$
12	$2B_2$	$\mathbb{Z}_8A$	$8A_1$
12	$2B_2$	$\mathbb{Z}_8A$	$8A_4$
12	$2A_4$	$\mathbb{Z}_8A$	$8A_1$
12	$2A_4$	$\mathbb{Z}_8A$	$8A_4$
12	$4A_1$	$\mathbb{Z}_8A$	$8A_1$
12	$4A_1$	$\mathbb{Z}_8A$	$8A_4$
12	$4A_2$	$\mathbb{Z}_8A$	$8A_1$
12	$4A_2$	$\mathbb{Z}_8A$	$8A_4$
17	$2A_2$	18	$2A_5$
17	$2A_2$	18	$4B_9$
17	$4B_8$	18	$2A_5$
17	$4B_8$	18	$4B_9$
18	$2A_5$	17	$2A_2$
18	$2A_5$	17	$4B_8$
18	$4B_9$	17	$2A_2$
18	$4B_9$	17	$4B_8$
19	$2A_3$	19	$2A_3$
19	$2A_3$	19	$4B_3$
19	$4B_3$	19	$2A_3$
19	$4B_3$	19	$4B_3$
20	$2B_3$	20	$2B_3$
20	$2B_3$	20	$2B_3$
20	$2B_3$	20	$4B_1$
20	$2B_3$	20	$4B_1$
20	$4B_1$	20	$2B_3$
20	$4B_1$	20	$2B_3$
20	$4B_1$	20	$4B_1$
20	$4B_1$	20	$4B_1$
21	$2B_1$	21	$2B_1$
21	$2B_1$	21	$4B_5$
21	$4B_5$	21	$2B_1$
21	$4B_5$	21	$4B_1$
23	$4A_3$	24	$4A_4$
23	$4A_3$	24	$4A_4$
23	$4A_3$	24	$4A_4$
23	$4A_3$	24	$4A_4$
24	$4A_4$	23	$4A_3$
24	$4A_4$	23	$4A_3$
24	$4A_4$	23	$4A_3$
24	$4A_4$	23	$4A_3$
25	$4B_4$	26	$4B_7$
25	$4B_4$	26	$4B_7$
25	$4B_4$	26	$4B_7$
25	$4B_4$	26	$4B_7$
26	$4B_7$	25	$4B_4$
26	$4B_7$	25	$4B_4$
26	$4B_7$	25	$4B_4$
26	$4B_7$	25	$4B_4$
27	$8A_2$	27	$8A_2$
27	$8A_2$	27	$8A_2$
27	$8A_2$	27	$8A_2$
27	$8A_2$	27	$8A_2$
27	$8A_3$	27	$8A_3$
27	$8A_3$	27	$8A_3$
27	$8A_3$	27	$8A_3$
27	$8A_3$	27	$8A_3$

Table 5: Relabeling for g in class 4B.

References

[1] T. Eguchi, H. Ooguri and Y. Tachikawa, "Notes on the K3 Surface and the Mathieu group M_{24}," Exper. Math. **20**, 91 (2011) `[arXiv:1004.0956 [hep-th]]`.

[2] M.C.N. Cheng, "K3 Surfaces, N=4 dyons, and the Mathieu group M_{24}," Commun. Number Theory Phys. **4**, 623 (2010) `[arXiv:1005.5415 [hep-th]]`.

[3] M.R. Gaberdiel, S. Hohenegger and R. Volpato, "Mathieu twining characters for K3," JHEP **1009**, 058 (2010) `[arXiv:1006.0221 [hep-th]]`.

[4] M.R. Gaberdiel, S. Hohenegger and R. Volpato, "Mathieu Moonshine in the elliptic genus of K3," JHEP **1010**, 062 (2010) `[arXiv:1008.3778 [hep-th]]`.

[5] T. Eguchi and K. Hikami, "Note on twisted elliptic genus of K3 surface," Phys. Lett. B **694**, 446 (2011) `[arXiv:1008.4924 [hep-th]]`.

[6] A. Taormina and K. Wendland, "The Symmetries of the tetrahedral Kummer surface in the Mathieu group M_{24}," arXiv:1008.0954 [hep-th].

[7] M.C.N. Cheng and J.F.R. Duncan, "On Rademacher sums, the largest Mathieu group, and the holographic modularity of Moonshine," Commun. Number Theory Phys. **6** (3), 697–758 (2012) `arXiv:1110.3859 [math.RT]`.

[8] M.R. Gaberdiel, S. Hohenegger and R. Volpato, "Symmetries of K3 sigma models," Commun.Num.Theor.Phys. **6** 1–50 (2012) `arXiv:1106.4315 [hep-th]`.

[9] A. Taormina and K. Wendland, "The overarching finite symmetry group of Kummer surfaces in the Mathieu group M_{24}," JHEP **1308** (2013) 125 `[arXiv:1107.3834 [hep-th]]`.

[10] R. Volpato, "Mathieu Moonshine and symmetries of K3 sigma models," Fortsch.Phys. **60**, 1112-1117 (2012) `[arXiv:1201.6172]`.

[11] M.R. Gaberdiel and R. Volpato, "Mathieu moonshine and orbifold K3s," in *Conformal Field Theory, Automorphic Forms and Related Topics*, Contrib. Math. Comput. Sci. **8**, Springer-Verlag, Berlin - Heidelberg - New York (2014) `[arXiv:1206.5143 [hep-th]]`.

[12] T. Gannon, "Much ado about Mathieu," arXiv:1211.5531 [math.RT].

[13] A. Taormina and K. Wendland, "Symmetry-surfing the moduli space of Kummer K3s," arXiv:1303.2931 [hep-th].

[14] A. Taormina and K. Wendland, "A twist in the M24 moonshine story," arXiv:1303.3221 [hep-th].

[15] M. R. Gaberdiel, A. Taormina, R. Volpato and K. Wendland, "A K3 sigma model with $\mathbb{Z}_2^8 : M_{20}$ symmetry," JHEP **1402** (2014) 022 `[arXiv:1309.4127 [hep-th]]`.

[16] T. Creutzig and G. Hoehn, "Mathieu Moonshine and the Geometry of K3 Surfaces," Commun. Num. TheorPhys. **08** (2014) 295 `[arXiv:1309.2671 [math.QA]]`.

[17] M. R. Gaberdiel, D. Persson, H. Ronellenfitsch and R. Volpato, "Generalized Mathieu Moonshine," Commun. Num. Theor. Phys. **7**, 145-223 (2013) `[arXiv:1211.7074 [hep-th]]`.

[18] M. R. Gaberdiel, D. Persson and R. Volpato, "Generalized Moonshine and Holomorphic Orbifolds," arXiv:1302.5425 [hep-th].

[19] S. Norton, "Generalized moonshine", Proc. Sympos. Pure Math. **47**, 209210, Amer. Math. Soc., Providence, RI (1987).

[20] M. C. N. Cheng, X. Dong, J. Duncan, J. Harvey, S. Kachru and T. Wrase, "Mathieu Moonshine and N=2 String Compactifications," JHEP **1309**, 030 (2013) `[arXiv:1306.4981 [hep-th]]`.

[21] S. Harrison, S. Kachru and N. M. Paquette, "Twining Genera of (0,4) Supersymmetric Sigma Models on K3," JHEP **1404** (2014) 048 `[arXiv:1309.0510 [hep-th]]`.

[22] J. A. Harvey and S. Murthy, "Moonshine in Fivebrane Spacetimes," JHEP **1401** (2014) 146 `[arXiv:1307.7717 [hep-th].]`

[23] R. Dijkgraaf, G.W. Moore, E.P. Verlinde and H.L. Verlinde, "Elliptic genera of symmetric products and second quantized strings," Commun. Math. Phys. **185**, 197 (1997) `[arXiv:hep-th/9608096]`.

[24] A. Sen, "A Twist in the Dyon Partition Function," JHEP **1005** (2010) 028 `[arXiv:0911.1563 [hep-th]]`.

[25] A. Sen, "Discrete Information from CHL Black Holes," JHEP **1011** (2010) 138 `[arXiv:1002.3857 [hep-th]]`.

[26] S. Govindarajan, "BKM Lie superalgebras from counting twisted CHL dyons," JHEP **1105** (2011) 089 `[arXiv:1006.3472 [hep-th]]`.

[27] S. Govindarajan, “Unravelling Mathieu Moonshine,” Nucl. Phys. B **864** (2012) 823 `[arXiv:1106.5715 [hep-th]]`.

[28] M.C.N. Cheng and J.F.R. Duncan, “The largest Mathieu group and (mock) automorphic forms,” Proc. Sympos. Pure Math. **85**, 53–82 (2012) `arXiv:1201.4140 [math.RT]`.

[29] M. Raum, “Powers of M24-twisted Siegel product expansions are modular,” `arXiv:1208.3453v2 [math.NT]`.

[30] M. C. N. Cheng, J. F. R. Duncan and J. A. Harvey, “Umbral Moonshine,” Commun. Num. TheorPhys. **08** (2014) 101 `[ arXiv:1204.2779 [math.RT]]`.

[31] G. Mason, “M_{24} and certain automorphic forms”, Contemp. Math. **45**, p. 223-244, (1985).

[32] G. Mason, “Elliptic systems and the eta-function,” Notas Soc. Mat. Chile **8**, 37 (1989).

[33] G. Mason, “On a system of elliptic modular forms attached to the largest Mathieu group”, Nagoya Math. J. **118**, 177 (1990).

[34] G. Mason, “G-elliptic systems and the genus zero problem for M_{24}”, Bull. Amer. Math. Soc. **25**, no. 1, p. 45-53, (1991).

[35] J.H. Conway and S.P. Norton, “Monstrous Moonshine”, Bull. London Math. Soc. **11**, 308 (1979).

[36] I. Frenkel, J. Lepowsky and A. Meurman, “Vertex operator algebras and the Monster”, Pure and Applied Mathematics **134**, Academic Press, Inc., Boston, MA, (1988).

[37] R.E. Borcherds, “Monstrous moonshine and monstrous Lie superalgebras” Invent. Math. **109**, 405 (1992).

[38] M.P. Tuite, “Generalized moonshine and Abelian orbifold constructions,” Contemp. Math. **193**, 353–368 (1996) `arXiv:hep-th/9412036`.

[39] C.-y. Dong, H.-s. Li and G. Mason, “Modular invariance of trace functions in orbifold theory,” Commun. Math. Phys. **214**, 1 (2000) `[arXiv:q-alg/9703016]`.

[40] G. Höhn, “Generalized moonshine for the baby Monster,” Habilitation thesis, Mathematisches Institut, Universität Freiburg (2003), `http://www.math.ksu.edu/ gerald/papers/baby8.ps`.

[41] S. Carnahan, "Generalized Moonshine I: Genus zero functions," Algebra Number Theory **4** (6), 649–679 (2010) `arXiv:0812.3440 [math.RT]`.

[42] S. Carnahan, "Generalized Moonshine II: Borcherds products," Duke Math. J. **161** (5), 893–950 (2012) `arXiv:0908.4223 [math.RT]`.

[43] S. Carnahan, "Generalized Moonshine III: Equivariant intertwining operators," to appear.

[44] S. Carnahan, "Generalized Moonshine IV: Monstrous Lie algebras," `arXiv:1208.6254 [math.RT]`.

[45] R. Dijkgraaf and E. Witten, "Topological gauge theories and group cohomology," Commun. Math. Phys. **129**, 393 (1990).

[46] P. Roche, V. Pasquier and R. Dijkgraaf, "Quasi-Hopf algebras, group cohomology and orbifold models," Nucl. Phys. Proc. Suppl. **18B**, 60 (1990).

[47] P. Bantay, "Orbifolds and Hopf algebras," Phys. Lett. B **245**, 477 (1990).

[48] A. Coste, T. Gannon and P. Ruelle, "Finite group modular data," Nucl. Phys. B **581**, 679 (2000) `[arXiv:hep-th/0001158]`.

[49] T. Gannon, talk at 'Mathieu Moonshine', ETH Zürich, July 2011.

[50] V. Gritsenko, "Elliptic genus of Calabi-Yau manifolds and Jacobi and Siegel modular forms," St. Petersburg Math. J. **11** (5), 781–804 (2000) `arXiv:math/9906190`.

[51] V. A. Gritsenko and V. V. Nikulin, "Automorphic Forms and Lorentzian Kac-Moody Algebras. II", Internat. J. Math. **9** (2), 201–275 (1998) `[arXiv:alg-geom/9611028]`.

[52] R. Dijkgraaf, E.P. Verlinde and H.L. Verlinde, "Counting dyons in N=4 string theory," Nucl. Phys. B **484**, 543 (1997) `[arXiv:hep-th/9607026]`.

[53] L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, "Strings on Orbifolds," Nucl. Phys. B **261** (1985) 678; "Strings on Orbifolds. 2.," Nucl. Phys. B **274** (1986) 285.

[54] L. Dolan, P. Goddard and P. Montague, "Conformal Field Theory of Twisted Vertex Operators," Nucl. Phys. B **338** (1990) 529.

[55] N. Ganter, "Hecke operators in equivariant elliptic cohomology and generalized moonshine", CRM Proc. Lecture Notes **47**, 173, Amer. Math. Soc., Providence, RI (2009) `[arXiv:0706.2898 [math.AT]]`.

[56] V. Ginzburg, M. Kapranov, E. Vasserot, "Elliptic algebras and equivariant elliptic cohomology I." [arXiv:q-alg/9505012]

[57] M. Eichler and D. Zagier, *The Theory of Jacobi Forms*, Birkhäuser (1985).

[58] R. Borcherds, "Automorphic forms on $\mathrm{O}_{s+2,2}(\mathbf{R})$ and infinite products," Invent. Math. **120** (1), 161 (1995).

[59] A. Dabholkar and S. Nampuri, "Spectrum of dyons and black holes in CHL orbifolds using Borcherds lift," JHEP **0711** (2007) 077 `[hep-th/0603066]`.

[60] V. A. Gritsenko and V. V. Nikulin, "The Igusa modular form and the "simplest" Lorentzian Kac-Moody algebras", Sb. Math. **187** (11), 1601–1641 (1996) `[arXiv:alg-geom/9603010]`.

[61] V. A. Gritenko and V. V. Nikulin, "Siegel automorphic form corrections of some Lorentzian Kac–Moody Lie algebras", Amer. J. Math. **119** (1), 181–224 (1997) `[arXiv:alg-geom/9504006]`.

[62] T. Eguchi and K. Hikami, "Twisted Elliptic Genus for K3 and Borcherds Product," Lett. Math. Phys. **102**, 203 (2012) `[arXiv:1112.5928 [hep-th]]`.

[63] F. Clery and V. A. Gritsenko, "The Siegel modular forms of genus 2 with the simplest divisor", Proc. Lond. Math. Soc. **102** (6), 1024–1052 (2011) `[arXiv:0812.3962 [math.NT]]`.

[64] J.R. David, D.P. Jatkar and A. Sen, "Product representation of Dyon partition function in CHL models," JHEP **0606**, 064 (2006) `[arXiv:hep-th/0602254]`.

[65] J.R. David, D.P. Jatkar and A. Sen, "Dyon spectrum in generic N=4 supersymmetric Z(N) orbifolds," JHEP **0701**, 016 (2007) `[arXiv:hep-th/0609109]`.

[66] H. Bass, J. Milnor and J. P. Serre, "Solution of the congruence subgroup problem for $\mathrm{SL}_n\,(n \geq 3)$ and $\mathrm{Sp}_{2n}\,(n \geq 2)$", Inst. Hautes Études Sci. Publ. Math. **33** (1967).

[67] M. C. N. Cheng, J. F. R. Duncan and J. A. Harvey, "Umbral Moonshine and the Niemeier Lattices," Res Math Sci 2014, **1**:3; [arXiv:1307.5793 [math.RT]].

[68] M.C.N. Cheng and E.P. Verlinde, "Wall crossing, discrete attractor flow, and Borcherds algebra," SIGMA **4**, 068 (2008) `[arXiv:0806.2337 [hep-th]]`.

[69] M. C. N. Cheng and A. Dabholkar, "Borcherds-Kac-Moody Symmetry of N=4 Dyons," Commun. Num. Theor. Phys. **3** (2009) 59 `[arXiv:0809.4258 [hep-th]]`.

[70] S. Govindarajan and K. Gopala Krishna, "Generalized Kac-Moody Algebras from CHL dyons," JHEP **0904** (2009) 032 `[arXiv:0807.4451 [hep-th]]`.

[71] S. Govindarajan and K. Gopala Krishna, "BKM Lie superalgebras from dyon spectra in Z(N) CHL orbifolds for composite N," JHEP **1005** (2010) 014 `[arXiv:0907.1410 [hep-th]]`.

[72] G. Karpilovsky, "Group Representations: Volume 2," North-Holland Mathematical Studies **177**, Elsevier (1993).

[73] M.D.F. de Wild Propitius, "Topological interactions in broken gauge theories," `arXiv:hep-th/9511195`.

arXiv:1211.7074v3 [hep-th] 16 Jan 2014

GENERALISED MATHIEU MOONSHINE

MATTHIAS R. GABERDIEL, DANIEL PERSSON, HENRIK RONELLENFITSCH,
AND ROBERTO VOLPATO

Abstract. The Mathieu twisted twining genera, i.e. the analogues of Norton's generalised Moonshine functions, are constructed for the elliptic genus of K3. It is shown that they satisfy the expected consistency conditions, and that their behaviour under modular transformations is controlled by a 3-cocycle in $H^3(M_{24}, U(1))$, just as for the case of holomorphic orbifolds. This suggests that a holomorphic VOA may be underlying Mathieu Moonshine.

Contents

1. Introduction 2
1.1. Monstrous Moonshine 2
1.2. Mathieu Moonshine 3
1.3. Generalised Monstrous Moonshine 4
1.4. Generalised Mathieu Moonshine 4
1.5. Outline 5
2. Twisted Twining Genera 6
2.1. Properties of the twisted twining genera 6
2.2. Cohomological obstructions 9
2.3. Classification of independent twisted twining genera 10
2.4. Modular properties of the twisted twining genera 12
2.5. Explicit twining genera 13
3. Holomorphic Orbifolds and Group Cohomology 15
3.1. Review of holomorphic orbifolds 16
3.2. Application to Mathieu Moonshine 18
3.3. Determining the cohomology class 20
3.4. Computation of the twisted twining characters: an example 24
3.5. Projective representations 27
4. K3 Orbifolds 27
4.1. The relabelling phenomenon 28

4.2. An example 30
4.3. Computation of a twisted twining genus 32
5. Conclusions 34
5.1. Summary 34
5.2. Open problems and future work 35
Acknowledgements 37
Appendix A. Details on the unobstructed twisted twining genera 37
A.1. The characters $\phi_{2\mathrm{B},4\mathrm{A}_2}$ (group 13) $\phi_{4\mathrm{B},4\mathrm{A}_3}$ (group 23) and $\phi_{4\mathrm{B},4\mathrm{A}_4}$ (group 24) 38
A.2. The cases $\phi_{3\mathrm{A},3\mathrm{A}_3}$ (group 33) and $\phi_{3\mathrm{A},3\mathrm{B}_1}$ (group 34) 40
Appendix B. Some group cohomology 42
Appendix C. Projective representations of finite groups 43
C.1. Central extension and orbifolds 45
Appendix D. Centralisers $C_{M_{24}}(g)$ and (projective) character tables 46
D.1. Character Tables 47
Appendix E. Decompositions of twisted sectors 61
References 68

1. INTRODUCTION

1.1. **Monstrous Moonshine.** In mathematics and physics the word ‘Moonshine’ has come to represent various surprising and deep connections between a priori unrelated fields, such as number theory, representation theory of finite groups, algebra and quantum field theory. The first and most well-known example of such a connection is, of course, Conway and Norton’s *Monstrous Moonshine conjecture* [1]. Their starting point was the observation of McKay that the first few Fourier coefficients of the modular invariant J-function $J(\tau)$ are dimensions of representations of the Monster group $\mathbb{M}$, the largest finite simple sporadic group. Conway and Norton then conjectured that to each element $g \in \mathbb{M}$ of the Monster, one can associate a function $T_g(\tau)$ (the so-called McKay-Thompson series) on the upper half-plane $\mathbb{H}_+$ that is invariant under some subgroup $\Gamma_g \subset SL(2,\mathbb{R})$. They also conjectured that the invariance group Γ_g must be genus zero, meaning that, as a Riemann surface, the quotient $\Gamma_g \backslash \mathbb{H}_+$ is topologically a sphere. Furthermore, they conjectured that $T_g(\tau)$ is in fact the Hauptmodul for Γ_g.

Subsequently, Frenkel, Lepowsky and Meurman (FLM) [2] constructed a graded $\mathbb{M}$-module $V^\natural = \bigoplus_{n=-1}^{\infty} V_n^\natural$, such that the dimension of the graded subspaces $a(n) = \dim V_n^\natural$ reproduce precisely the Fourier coefficients in the q-expansion of the J-function, $J(\tau) = \sum_{n=-1}^{\infty} a(n)q^n$, $q = e^{2\pi i\tau}$. The coefficients $a_g(n)$ in the Fourier expansion of the McKay-Thompson series are then identified with the characters $\mathrm{Tr}_{V_n^\natural}(g)$. Physically, the FLM-module is the $\mathbb{Z}_2$-orbifold of the open bosonic string compactified on the Leech torus $(\mathbb{R}^{24}/\Lambda_{\text{Leech}})$, where Λ_{Leech} is the Leech lattice. The FLM-construction provided important clues into the moonshine conjectures, but fell short of proving them. The final proof was found by Borcherds [3], making heavy use of the FLM-module $V^\natural$, but also introducing yet another set of ingredients to the story, namely generalised Kac-Moody algebras and automorphic denominator identities. Thus, by the time the original conjectures were proven, Monstrous Moonshine encompassed not only the realms of finite groups and modular forms, but also string theory and infinite-dimensional algebras, for a nice review see [4].

1.2. **Mathieu Moonshine.**

Recently, a very interesting new moonshine phenomenon was conjectured by Eguchi, Ooguri and Tachikawa (EOT) [5]: they observed that the first few Fourier coefficients of the elliptic genus of K3 are dimensions of representations of the largest Mathieu group M_{24}. This suggests the existence of a new moonshine-type connection between Mathieu groups, Jacobi forms and K3 surfaces. The analogues of the McKay-Thompson series, the *twining elliptic genera* $\phi_g(\tau, z)$, $g \in M_{24}$, were constructed in a series of papers [6, 7, 8, 9], and it was shown that they are all weak Jacobi forms of weight 0 and index 1 (with multiplier system) for subgroups $\Gamma_0(N_g) \subset SL(2, \mathbb{Z})$, where N_g is the order of the element g. (The elliptic genus ϕ_{K3} itself corresponds to taking g to be the identity element.) The compatibility with M_{24}-representations was checked up to the first 600 coefficients [8, 9]; according to Gannon [10] this is sufficient to prove that all Fourier coefficients of the elliptic genus of K3 are dimensions of M_{24} representations. It is also shown in [10] that the multiplicities of the real M_{24} representations are always even, see also [11].

Although in a certain sense the above results establish the Mathieu Moonshine conjecture, there are many aspects of it that are much less understood compared to Monstrous Moonshine. For instance, the genus zero property of Monstrous Moonshine does not seem to hold for the M_{24}-twining genera since some of the invariance groups $\Gamma_0(N_g)$ are not genus zero.[*] More importantly, the analogue of the FLM module $V^\natural$ is not yet known for Mathieu Moonshine, i.e. we do not know of any CFT with automorphism group M_{24} whose elliptic genus reproduces the elliptic genus of K3. In particular, none of the K3-sigma models have this property since M_{24} is not contained in the automorphism group of any of them [13].

---

[*] It was proposed in [12] that the correct generalisation of the genus zero property is a certain Rademacher-summability condition, which is indeed satisfied in all cases.

1.3. **Generalised Monstrous Moonshine.** A few years after the original Monstrous Moonshine conjectures, Norton proposed a generalisation that he dubbed *Generalised Moonshine* [14]. He suggested that for each commuting pair of Monster group elements $g, h \in \mathbb{M}$ there exists a function $f(g, h; \tau)$ that is also invariant under a genus zero subgroup of $SL(2, \mathbb{R})$. These functions generalise the McKay-Thompson series to which they reduce for $g = e$. Furthermore, they transform into one another under a simultaneous action of $SL(2, \mathbb{Z})$ on $(\tau; g, h) \in \mathbb{H}_+ \times \mathcal{P}(\mathbb{M})$, where $\mathcal{P}(\mathbb{M}) \subset \mathbb{M} \times \mathbb{M}$ is the set of pairs of commuting elements of $\mathbb{M}$. Finally, the Fourier coefficients in the q-expansion of $f(g, h; \tau)$ are conjectured to be dimensions of (projective) representations of the centraliser of g in $\mathbb{M}$, $C_{\mathbb{M}}(g) = \{k \in \mathbb{M} : gk = kg\}$. Although this conjecture has been proven in special cases [15, 16, 17], the general case is still open.[†]

The Generalised Moonshine conjecture was given a physical interpretation by Dixon, Ginsparg and Harvey [24] that was later elaborated upon by Ivanov and Tuite [17, 18, 19]. They showed that the Norton series $f(g, h; \tau)$ arises naturally as the character in the twisted sector $V^\natural_g$ of an orbifold of the Monster CFT $V^\natural$ by the element $g \in \mathbb{M}$, 'twined' by the group element h; in standard CFT language, they can therefore be interpreted as

$$f(g,h;\tau) = g\, \underset{h}{\boxed{}} = \mathrm{Tr}_{V^\natural_g}\left(h\, q^{L_0-1}\right) \ . \tag{1}$$

Many of the properties conjectured by Norton can be proved from holomorphic orbifold considerations [15].

1.4. **Generalised Mathieu Moonshine.** In this paper, we show that Norton's generalisation also applies to Mathieu Moonshine. More specifically, for every pair of commuting group elements $g, h \in M_{24}$, we construct 'twisted twining genera' $\phi_{g,h} : \mathbb{H}_+ \times \mathbb{C} \to \mathbb{C}$ that either vanish or are weak Jacobi forms of weight 0 and index 1 for some $\Gamma_{g,h} \subset SL(2, \mathbb{Z})$. For $g = e$, they reduce to the twining genera of [6, 7, 8, 9], and they transform under the modular group $SL(2, \mathbb{Z})$ into one another. The multiplier phases that appear in these transformations behave as though these twisted twining genera were twisted twining characters of a holomorphic orbifold; in particular, they are controlled by a 3-cocycle $\alpha \in H^3(M_{24}, U(1))$ via a formula that was first written down by Dijkgraaf and Witten in [25]. Furthermore, the Fourier coefficients of the twisted elliptic genera $f(g, e; \tau)$ equal dimensions of projective representations of the centraliser $C_{M_{24}}(g)$, and the central extension that characterises the projective representation is again determined by the cohomology class $\alpha \in H^3(M_{24}, U(1))$.

[†] Carnahan has announced a series of papers [20, 21, 22, 23] which he claims will lead to a complete proof.

The idea for using the cohomology group $H^3(M_{24}, U(1))$ and the formalism of Dijkgraaf and Witten in order to understand the multiplier phases of the twisted twining genera was first suggested to us by Terry Gannon in 2011 [26], see also [10]. Mason has also speculated [27] that $H^3(\mathbb{M}, U(1))$ should play a similar role in the context of Generalised Monstrous Moonshine, although this has, to our knowledge, not yet been worked out (partially, because $H^3(\mathbb{M}, U(1))$ is unknown).

We should also mention that Mason has previously proposed a version of Generalised Moonshine for M_{24} [28], where the role of the Norton series $f(g, h; \tau)$ is played by products of eta functions $\eta(\tau)$. The precise relation between our twisted twining genera and Mason's eta products will be explained elsewhere [29].

1.5. **Outline.** The paper is organised as follows. In section 2 we introduce the twisted twining genera $\phi_{g,h}$, and discuss their general properties, in particular the expected behaviour under modular transformations. Based on these properties we then explain how many independent twisted twining genera there are, and list all of them explicitly, see table 3. The detailed derivation of these genera is illustrated in section 3. First, we review the structure of holomorphic orbifolds and explain in particular, the role of the 3-cocycle in characterising the various transformation properties of the twisted twining characters. We then postulate that the transformation properties of the twisted twining genera are similarly constrained. In support of this assumption, we prove that there exists a unique 3-cocycle in $H^3(M_{24}, U(1))$ compatible with the known properties of the (untwisted) twining genera of [6, 7, 8, 9]. This then allows us to find all twisted twining genera explicitly. We also check that, up to the first 500 levels in each twisted sector, the resulting functions are compatible with the requirement that they arise from the appropriate projective representation of the centraliser. In section 4 we subject our results to two independent consistency checks. First, for group elements whose orbifold leads again to a K3 sigma-model, we calculate the twining genera of the orbifold from the twisted twining genera of the original theory, and show that we reproduce answers from [8, 9]. However, as it turns out, the relevant group element is sometimes different from what one would have expected, and we explain this 'relabelling' phenomenon in quite some detail in section 4.1, see also section 4.2 for an explicit example. Secondly, we explain the vanishing of some of the twisted twining genera from a geometrical point of view, see section 4.3. Finally we end with some conclusions and open problems in section 5. We have relegated some of the more technical material to various appendices to which we refer throughout the main body of the paper. Moreover, the appendices contain the character tables of the centralisers and the decompositions of the twisted sectors, up to the first 20 levels. The decompositions up to 500 levels, as well as the details of the computations requiring computer support, are collected in the ancillary files in the arXiv repository of this paper.

2. Twisted Twining Genera

Let us begin by introducing the twisted twining genera that are the main object of this paper. Suppose we are given a K3 sigma-model $\mathcal{H}$, i.e. a CFT describing string propagation on K3, whose automorphism group contains two commuting elements $g, h \in M_{24}$. Then we can consider the orbifold of the sigma-model by g, and in particular, define the g-twisted sector $\mathcal{H}_g$. Since g and h commute, h gives rise to an action on $\mathcal{H}_g$, and we can define the *twisted twining genus* $\phi_{g,h}(\tau, z)$ by

$$\phi_{g,h}(\tau,z) = \operatorname{Tr}_{\mathcal{H}_g}\Big(h(-1)^{J_0+\bar{J}_0} q^{L_0-\frac{c}{24}} \bar{q}^{\bar{L}_0-\frac{\bar{c}}{24}} y^{J_0}\Big)\ , \tag{2}$$

where $q = e^{2\pi i\tau}$ and $y = e^{2\pi i z}$. We expect that $\phi_{g,h} : \mathbb{H}_+ \times \mathbb{C} \to \mathbb{C}$ is holomorphic in both τ and z. Furthermore, since the elliptic genus is independent of the choice of the underlying K3 sigma-model, we expect that the same is true for these twisted twining genera, i.e. we expect that (2) does not depend on the choice of $\mathcal{H}$ (as long as g and h are automorphisms of $\mathcal{H}$). By construction, for $g = e$ the identity element in M_{24}, the twisted twining genus $\phi_{e,h}$ agrees with the twining genus ϕ_h considered in [6, 7, 8, 9]; in particular, for $g = h = e$ $\phi_{e,e}$ is just the elliptic genus of K3.

Unfortunately, while for some commuting pairs (g, h) an actual K3 sigma-model for which both g and h are automorphisms can be found, this is not true in general [13]; in fact, this problem already arises for the usual twining genera, i.e. for the pairs (e, h). In the spirit of the EOT conjecture, we shall nevertheless assume that twisted twining genera $\phi_{g,h}$ can be defined — albeit not directly by a formula of the form (2) — for *all* commuting pairs $g, h \in M_{24}$. The fact that our construction will be successful is an a posteriori justification for this assumption.

2.1. **Properties of the twisted twining genera.** The definition of the twisted twining genera $\phi_{g,h}$ in terms of (2) suggests that they should satisfy the following properties:

(A) Elliptic and modular properties:

$$\phi_{g,h}(\tau, z+\ell\tau+\ell') = e^{-2\pi i(\ell^2\tau+2\ell z)}\,\phi_{g,h}(\tau,z) \qquad \ell,\ell'\in\mathbb{Z} \tag{3}$$

$$\phi_{g,h}\Big(\frac{a\tau+b}{c\tau+d},\frac{z}{c\tau+d}\Big) = \chi_{g,h}\big(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\big)\, e^{2\pi i\frac{cz^2}{c\tau+d}}\,\phi_{h^cg^a,h^dg^b}(\tau,z)\ , \quad \big(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\big)\in SL(2,\mathbb{Z}) \tag{4}$$

for a certain multiplier $\chi_{g,h} : SL(2,\mathbb{Z}) \to U(1)$. In particular, each $\phi_{g,h}$ is a weak Jacobi form of weight 0 and index 1 with multiplier $\chi_{g,h}$ under a subgroup $\Gamma_{g,h}$ of $SL(2,\mathbb{Z})$ (see [30] for the definitions).

(B) Invariance under conjugation of the pair g, h in M_{24},

$$\phi_{g,h}(\tau,z) = \xi_{g,h}(k)\,\phi_{k^{-1}gk,k^{-1}hk}(\tau,z)\ , \qquad k\in M_{24}\ , \tag{5}$$

where $\xi_{g,h}(k)$ is a phase.

(C) If $g\in M_{24}$ has order N, the twisted twining genera $\phi_{g,h}$ have an expansion of the form

$$\phi_{g,h}(\tau,z) = \sum_{\substack{r\in\lambda_g+\mathbb{Z}/N\\ r\geq 0}} \mathrm{Tr}_{\,\mathcal{H}_{g,r}}\big(\rho_{g,r}(h)\big)\,\mathrm{ch}_{h=\frac{1}{4}+r,\ell}(\tau,z)\ , \tag{6}$$

where $\lambda_g\in\mathbb{Q}$, and $\mathrm{ch}_{h,\ell}(\tau,z)$ are elliptic genera of Ramond representations of the $\mathcal{N}=4$ superconformal algebra at central charge $c=6$. (Here $\ell=\frac{1}{2}$, except possibly for $h=\frac{1}{4}$, where $\ell=0$ is also possible — if both $\ell=0,\frac{1}{2}$ appear for $r=0$, it is understood that there are two such terms in the above sum. We use the same conventions for the elliptic genera as e.g. in [31].) Furthermore, each vector space $\mathcal{H}_{g,r}$ is finite dimensional, and it carries a projective representation $\rho_{g,r}$ of the centraliser $C_{M_{24}}(g)$ of g in M_{24}, such that

$$\rho_{g,r}(g) = e^{2\pi i r}\ , \qquad \rho_{g,r}(h_1)\,\rho_{g,r}(h_2) = c_g(h_1,h_2)\,\rho_{g,r}(h_1h_2)\ , \tag{7}$$

for all $h_1,h_2\in C_{M_{24}}(g)$. Here $c_g : C_{M_{24}}(g)\times C_{M_{24}}(g)\to U(1)$ is independent of r, and satisfies the cocycle condition

$$c_g(h_1,h_2)\,c_g(h_1h_2,h_3) = c_g(h_1,h_2h_3)\,c_g(h_2,h_3) \tag{8}$$

for all $h_1,h_2,h_3\in C_{M_{24}}(g)$.

(D) For $g=e$, where e is the identity element of M_{24}, the functions $\phi_{e,h}$ correspond to the twining genera considered in [6]–[9]. In particular, $\phi_{e,e}$ is the K3 elliptic genus.

Since the elliptic genus behaves essentially like the character of a holomorphic CFT — in particular, it is modular invariant and holomorphic by itself — it is natural to believe that the same will be true for the twisted twining genera, i.e. that they will be analogous to twisted twining characters of a holomorphic CFT. As we will review in more detail in section 3, the modular properties of the twisted twining characters of a holomorphic CFT are controlled by a 3-cocycle $\alpha : G \times G \times G \to U(1)$ representing a cohomology class in the third cohomology group $H^3(G, U(1))$. (Some background material about group cohomology can be found in appendix B.) We will therefore postulate that

(E) The multipliers $\chi_{g,h}$, the phases $\xi_{g,h}$, and the 2-cocycles c_g associated with the projective representations $\rho_{g,r}$ are completely determined (by the same formulas as for holomorphic orbifolds) in terms of a 3-cocycle α representing a class in $H^3(M_{24}, U(1))$.

The third cohomology group of M_{24} was only recently computed with the result [32][‡]

$$H^3(M_{24}, U(1)) \cong \mathbb{Z}_{12} \ . \tag{9}$$

The fact that this group is known explicitly plays a crucial role in our analysis. The specific cohomology class $[\alpha] \in H^3(M_{24}, U(1))$ that is relevant in our context is uniquely determined by the condition that it reproduces the multiplier system for the twining genera $\phi_{e,h}$ as described in [8], namely

$$\chi_{e,h}\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = e^{\frac{2\pi i cd}{o(h)\ell(h)}} \ , \qquad \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(o(h)) \ . \tag{10}$$

Here, $o(h)$ is the order of h and $\ell(h)$ is the length of the smallest cycle, when $h \in M_{24}$ is regarded as a permutation of 24 symbols [12]. Indeed, since $\ell(12\mathrm{B}) = 12$, it follows that α must correspond to the primitive generator of $H^3(M_{24}, U(1))$. The main result of our paper can now be stated as follows:

There exists a unique set of functions $\phi_{g,h}$ (unique up to redefinitions by (g,h)-dependent but otherwise constant phases) and a unique cohomology class $[\alpha] \in H^3(M_{24}, U(1))$ such that all conditions (A)–(E) are satisfied (condition (C) has been verified only for the first 500 representations $\rho_{g,r}$ for each $g \in M_{24}$).

When $g = e$ is the trivial element of M_{24}, the existence of all representations $\rho_{e,r}$ fitting eq. (6) has been recently proven in [10]. When $g \neq e$, condition (C) is, strictly speaking, still a conjecture. However, we provide compelling numerical evidence in favour of it, by

---

[‡] Note that for a finite group G one has the isomorphisms

$$H_{n-1}(G, \mathbb{Z}) \cong H^n(G, \mathbb{Z}), \qquad H^n(G, \mathbb{Z}) \cong H^{n-1}(G, U(1)) \ ,$$

which in particular imply that $H_3(M_{24}, \mathbb{Z}) \cong H^3(M_{24}, U(1))$.

verifying eq. (6) for the first 500 representations $\rho_{g,r}$ in each twisted sector. A complete proof of (6) should be possible along the lines of [10].

In the following we want to illustrate how the set of functions $\phi_{g,h}$ can be determined, and which explicit form it takes. As it turns out, a surprisingly large number of the twisted twining genera vanish, as we shall now explain.

2.2. **Cohomological obstructions.** For some pairs of commuting elements $g, h \in M_{24}$, the transformation properties above can only be satisfied if $\phi_{g,h}$ vanishes identically. In this case, we will say that the corresponding twisted twining genus is *obstructed*. As will be shown in more detail in section 3, these obstructions are also controlled by the cohomology class of the cocycle $\alpha \in H^3(M_{24}, U(1))$.[§]

In order to understand the origin of these obstructions, let us first derive some general consequences of our assumptions (A)–(E). First we note that the $SL(2,\mathbb{Z})$ action on the twisted twining characters can be extended to $GL(2,\mathbb{Z})$ by setting

$$\phi^*_{g,h}(\tau,z) = \chi_{g,h}\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)\, \phi_{g,h^{-1}}(\tau,z) \ , \tag{11}$$

where $\phi^*_{g,h}(\tau,z)$ is obtained by taking the complex conjugate of all Fourier coefficients of $\phi_{g,h}(\tau,z)$

$$\phi^*_{g,h}(\tau,z) = \overline{\phi_{g,h}(-\bar\tau,-\bar z)} \ . \tag{12}$$

The identity (11) follows from eq. (6), together with the observation that, for any projective representation R of a finite group G, $\mathrm{Tr}_R(h^{-1})$ equals $\overline{\mathrm{Tr}_R(h)}$ up to a phase, which we have denoted by $\chi_{g,h}\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$. As we will see, this phase also depends on the 3-cocycle α. We also note that since the $\mathcal{N}=4$ characters are invariant under $z \to -z$, we have the identity

$$\phi_{g,h}(\tau,-z) = \phi_{g,h}(\tau,z) \ . \tag{13}$$

With these preparations we can now describe two possible kinds of obstructions.

Obstruction 1: Let us consider pairwise commuting g, h, k. Then, by (5), we have

$$\phi_{g,h}(\tau,z) = \xi_{g,h}(k)\, \phi_{g,h}(\tau,z) \ , \tag{14}$$

and if

$$\xi_{g,h}(k) \neq 1 \ , \tag{15}$$

it follows that $\phi_{g,h}(\tau,z) = 0$.

[§] The fact that such obstructions might exist was first suggested to us by T. Gannon.

Obstruction 2: Suppose there are $g, h, k \in M_{24}$, with g and h commuting and

$$k^{-1}g^{-1}k = g \ , \qquad k^{-1}h^{-1}k = h \ , \tag{16}$$

i.e. the commuting pair (g, h) is conjugate within M_{24} to the pair (g^{-1}, h^{-1}). Then, by eqs. (51) and (42) (see section 3 for details), we obtain the relations

$$\phi_{g,h}(\tau, z) = \chi_{g,h}\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \right) \phi_{g^{-1},h^{-1}}(\tau, -z) \tag{17}$$

$$= \chi_{g,h}\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \right) \xi_{g^{-1},h^{-1}}(k)\, \phi_{g,h}(\tau, -z) \ . \tag{18}$$

Therefore, if

$$\chi_{g,h}\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix} \right) \xi_{g^{-1},h^{-1}}(k) \neq 1 \ , \tag{19}$$

eq. (13) implies $\phi_{g,h}(\tau, z) = 0$.
As we shall see, these two obstructions are responsible for the fact that most twisted twining genera vanish.

2.3. Classification of independent twisted twining genera.

Our next aim is to enumerate all possible independent twisted twining genera. Let us denote the set of commuting group elements by

$$\mathcal{P} = \{(g, h) \in M_{24} \times M_{24} \mid gh = hg\} \ . \tag{20}$$

This set carries an action of $GL(2, \mathbb{Z}) \times M_{24}$, given by

$$(g, h) \mapsto (k^{-1}(g^a h^c)k, k^{-1}(g^b h^d)k) \ , \quad \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in GL(2, \mathbb{Z}) \ , \quad k \in M_{24} \ , \tag{21}$$

and the twisted twining genera associated to different (g, h)'s in the same orbit are related to one another by modular transformations, see eqs. (4), (5), and/or by complex conjugation, see eq. (11). We therefore want to describe the set of orbits of $\mathcal{P}$ under the action (21).

First we note that the $GL(2, \mathbb{Z})$ orbit of a pair (g, h) consists of all possible pairs of generators of the abelian group $\langle g, h \rangle \subset M_{24}$. Thus the orbits of $GL(2, \mathbb{Z}) \times M_{24}$ are in one-to-one correspondence with the conjugacy classes of abelian subgroups of M_{24} generated by two elements, i.e.

$$\bar{\mathcal{P}} = \mathcal{P}/(GL(2, \mathbb{Z}) \times M_{24}) = \{M_{24} - \text{conjugacy classes of groups } \langle g, h \rangle \subset M_{24}, gh = hg\} \ . \tag{22}$$

This description allows a complete classification of the orbits $[g, h] \in \bar{\mathcal{P}}$: there are 55 such orbits, 21 of which correspond to cyclic subgroups, i.e. to subgroups of the form $[e, h]$. The associated twisted twining genera are therefore just the twining genera $\phi_{e,h}$, for which explicit expressions were already derived in [8, 9]. Thus we only need to construct

#	Structure	Elements	$\lvert C(g,h)\rvert$	$\frac{\lvert N(g,h)\rvert}{\lvert C(g,h)\rvert}$	Orbits on **24**	Max subgr.
1.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	1536	6	$1^8 \cdot 4^4$	
2.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	1536	6	2^{12}	
3.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^3$	128	6	$1^4 \cdot 2^6 \cdot 4^2$	
4.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	3840	6	4^6	
5.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	96	6	4^6	
6.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{B})^3$	64	6	4^6	
7.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})^2(2\mathrm{B})$	256	2	$2^8 \cdot 4^2$	
8.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})(2\mathrm{B})^2$	512	2	$2^4 \cdot 4^4$	
9.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(2\mathrm{A})(2\mathrm{B})^2$	128	2	$2^4 \cdot 4^4$	
10.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^4$	64	8	$2^4 \cdot 8^2$	1
11.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^4$	64	8	4^6	2
12.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^2(2\mathrm{B})(4\mathrm{A})^2(4\mathrm{B})^2$	32	2	$2^2 \cdot 4^3 \cdot 8^1$	7
13.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{A})^4$	64	8	$4^2 \cdot 8^2$	8
14.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{A})^4$	32	8	$4^2 \cdot 8^2$	8
15.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{C})^4$	32	4	$4^2 \cdot 8^2$	8
16.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{C})^4$	16	4	$4^2 \cdot 8^2$	9
17.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	64	8	$1^4 \cdot 2^2 \cdot 8^2$	1
18.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	64	8	$2^4 \cdot 4^4$	2
19.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^4$	16	8	$1^2 \cdot 2^3 \cdot 4^2 \cdot 8^1$	3
20.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{B})^4$	64	8	$2^4 \cdot 8^2$	8
21.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{B})^4$	16	8	$2^4 \cdot 8^2$	9
22.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{A})^4(4\mathrm{C})^8$	16	16	$8^1 \cdot 16^1$	$13, 15, 15$
23.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^8(4\mathrm{B})^4$	16	32	$2^2 \cdot 4^1 \cdot 16^1$	$10, 10, 17$
24.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{A})^8(4\mathrm{B})^4$	16	32	$4^2 \cdot 8^2$	$11, 11, 18$
25.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^{12}$	16	96	$1^4 \cdot 4^1 \cdot 16^1$	$17, 17, 17$
26.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$(2\mathrm{A})^3(4\mathrm{B})^{12}$	16	96	4^6	$18, 18, 18$
27.	$\mathbb{Z}_2 \times \mathbb{Z}_8$	$(2\mathrm{A})(2\mathrm{B})^2(4\mathrm{B})^4(8\mathrm{A})^8$	16	8	$2^2 \cdot 4^1 \cdot 16^1$	20
28.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2\mathrm{A})^3(3\mathrm{A})^2(6\mathrm{A})^6$	12	12	$1^2 \cdot 3^2 \cdot 4^1 \cdot 12^1$	1
29.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2\mathrm{A})^3(3\mathrm{A})^2(6\mathrm{A})^6$	12	12	$2^3 \cdot 6^3$	2
30.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2\mathrm{B})^3(3\mathrm{B})^2(6\mathrm{B})^6$	12	12	12^2	4
31.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$(2\mathrm{B})^3(3\mathrm{B})^2(6\mathrm{B})^6$	12	12	12^2	5
32.	$\mathbb{Z}_2 \times \mathbb{Z}_{10}$	$(2\mathrm{B})^3(5\mathrm{A})^4(10\mathrm{A})^{12}$	20	12	$4^1 \cdot 20^1$	4
33.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$(3\mathrm{A})^8$	9	48	$1^3 \cdot 3^4 \cdot 9^1$	
34.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$(3\mathrm{A})^2(3\mathrm{B})^6$	9	12	$3^2 \cdot 9^2$	

TABLE 1. The 34 conjugacy classes of abelian subgroups of rank 2 in M_{24}.

the remaining 34 genuinely twisted twining genera. We have tabulated the corresponding conjugacy classes of groups $\langle g,h\rangle$ in table 1. There we have described their structure as an abelian group, i.e. as $\mathbb{Z}_m \times \mathbb{Z}_n$, the M_{24} classes of all its elements (excluding the identity), the order of the centraliser $C(g,h)$, the index $|N(g,h)|/|C(g,h)|$ of the centraliser in the normaliser of $\langle g,h\rangle$ in M_{24}, and the lengths of the orbits of $\langle g,h\rangle \subset M_{24}$ when acting as a group of permutations of 24 objects. Finally, the last column gives the conjugacy classes of the non-cyclic maximal subgroups. For example, group 24 has three distinct non-cyclic maximal subgroups, all of the form $\mathbb{Z}_2 \times \mathbb{Z}_4$; two of them are conjugated to group 11 and one is conjugated to group 18.

For M_{24} it turns out that the orbits of $\mathcal{P}$ under $SL(2,\mathbb{Z}) \times M_{24}$ are exactly the same as those under $GL(2,\mathbb{Z}) \times M_{24}$, i.e.[¶]

$$\bar{\mathcal{P}} = \mathcal{P}/(GL(2,\mathbb{Z}) \times M_{24}) = \mathcal{P}/(SL(2,\mathbb{Z}) \times M_{24}) \ . \tag{23}$$

Thus the twisted twining genera in each orbit in $\bar{\mathcal{P}}$ are just related by phases, see eqs. (4) and (5), and we do not need to invoke eq. (11).

2.4. **Modular properties of the twisted twining genera.** For each commuting pair $(g,h) \in \mathcal{P}$, let us denote by $\tilde{\Gamma}_{g,h} \subset SL(2,\mathbb{Z}) \times M_{24}$ the group of elements (γ, k) that leave the pair (g,h) fixed or map it to its inverse (g^{-1},h^{-1})

$$\tilde{\Gamma}_{g,h} = \left\{ \left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right), k \right) \in SL(2,\mathbb{Z}) \times M_{24} \mid \left(k^{-1}(g^a h^c)k,\, k^{-1}(g^b h^d)k \right) = (g,h) \text{ or } (g^{-1},h^{-1}) \right\} \ . \tag{24}$$

It follows from (4) and (5), together with (13) as well as (51) below, that the corresponding twisted twining genus $\phi_{g,h}$ will be invariant (up to a phase) under $\Gamma_{g,h} \equiv \pi(\tilde{\Gamma}_{g,h})$, where π denotes the projection of $SL(2,\mathbb{Z}) \times M_{24}$ onto its first factor, i.e. onto $SL(2,\mathbb{Z})$.

Each twisted twining genus $\phi_{g,h}$ belongs to a (finite-dimensional) vector representation of $SL(2,\mathbb{Z})$, that is spanned by the functions $\{\phi_{(g,h)\gamma}(\tau,z)\}$, where $\gamma \in SL(2,\mathbb{Z})$. Since every $\gamma \in \Gamma_{g,h}$ acts trivially (i.e. up to a phase) on $\phi_{g,h}$, the $SL(2,\mathbb{Z})$ representation is spanned by $\{\phi_{(g,h)\gamma}(\tau,z)\}_{\gamma \in \Gamma_{g,h}\backslash SL(2,\mathbb{Z})}$, where γ runs over a set of representatives for the cosets in $\Gamma_{g,h}\backslash SL(2,\mathbb{Z})$. Each of these functions $\phi_{(g,h)\gamma}$ is a weak Jacobi form of weight 0 and index 1 for the congruence subgroup $\gamma^{-1}\Gamma_{g,h}\gamma \subseteq SL(2,\mathbb{Z})$. We have tabulated for each orbit $[g,h] \in \bar{\mathcal{P}}$ the functions $\phi_{(g,h)\gamma}$ (starting from $\phi_{g,h}$) and the invariance group $\Gamma_{g,h}$ in table 2. (Note, however, that some of these functions vanish identically; in particular, this will be the case if there is an obstruction.) Each $\phi_{g,h}$ is denoted by the M_{24} class of g and the $C_{M_{24}}(g)$-class of h, named as in appendix D.

[¶] For a generic group G different from M_{24}, some $SL(2,\mathbb{Z}) \times G$ orbits might be strictly smaller than the $GL(2,\mathbb{Z}) \times G$ orbits.

Group 27 is the only case where the pairs (g, h) and (g^{-1}, h^{-1}) are not conjugated within M_{24}. Thus, charge conjugation gives the identities $\phi_{\mathrm{2B},\mathrm{8A}_1} = \phi_{\mathrm{2B},\mathrm{8A}_2}$, $\phi_{\mathrm{8A},\mathrm{2B}_1} = \phi_{\mathrm{8A},\mathrm{2B}_2}$, and so on, and the respective functions are denoted in table 2 by $\phi_{\mathrm{2B},\mathrm{8A}_{1,2}}$, $\phi_{\mathrm{8A},\mathrm{2B}_{1,2}}$, etcetera. Because of (11), most of the functions satisfy a 'reality' condition of the form

$$\phi^*_{g,h} = \zeta\, \phi_{g,h} \ , \tag{25}$$

for some constant phase ζ, which implies that the Fourier coefficients of $\zeta^{1/2}\phi_{g,h}$ are all real. Note that even if $\phi^*_{g,h}$ and $\phi_{g,h}$ are not proportional to one another, they are necessarily related by a modular transformation because of (23); if this is the case, they are listed as distinct functions in the same orbit in $\bar{\mathcal{P}}$.

Most of the modular groups $\Gamma_{g,h}$ are of the form $\Gamma(1) = SL(2,\mathbb{Z})$ or

$$\Gamma_0(N) := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid c \equiv 0 \mod N\} \ , \tag{26}$$

or conjugates of $\Gamma_0(N)$ in $SL(2,\mathbb{Z})$. The exceptions are the group in case 32, where

$$\Gamma_{\mathrm{2B},\mathrm{10A}} = \bigcup_{i\in\mathbb{Z}/3\mathbb{Z}, j\in\mathbb{Z}/4\mathbb{Z}} \left(\begin{smallmatrix} 1 & 1 \\ -5 & -4 \end{smallmatrix}\right)^i \left(\begin{smallmatrix} -3 & -1 \\ 10 & 3 \end{smallmatrix}\right)^j \Gamma_{2,10} \ , \tag{27}$$

is a subgroup of index 12 in $SL(2,\mathbb{Z})$ and

$$\Gamma_{2,10} := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid a \equiv 1,\ b \equiv 0 \mod 2,\ c \equiv 0,\ d \equiv 1 \mod 10\} \ , \tag{28}$$

is the group of elements $\gamma \in SL(2,\mathbb{Z})$ such that $(g, h)\cdot\gamma = (g, h)$; the group in case 12, with

$$\Gamma_{\mathrm{2A},\mathrm{4A}} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid b \equiv 0 \mod 2, c \equiv 0 \mod 4\} \ , \tag{29}$$

which is a conjugate of $\Gamma_0(8)$ in $SL(2,\mathbb{R})$; and the group in case 22, with

$$\Gamma_{\mathrm{4A},\mathrm{4C}} = \langle \left(\begin{smallmatrix} -1 & 1 \\ -2 & 1 \end{smallmatrix}\right) \ , \ \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right) \ , \ \left(\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right) \ , \ \left(\begin{smallmatrix} 3 & -2 \\ -4 & 3 \end{smallmatrix}\right) \rangle \ . \tag{30}$$

2.5. **Explicit twining genera.** We have now reduced our problem to finding 34 functions $\phi_{g,h}$ that are weak Jacobi forms with respect to $\Gamma_{g,h}$ up to some multiplier phases. As will be explained in the following section 3, our assumption (E) allows us to determine the precise form of these multipliers explicitly. Then most of the 34 functions vanish because of the obstructions, see table 3. For those that do not, the modular properties are strong enough to determine them explicitly, see also table 3. These explicit results are one of the main results of this paper, and their derivation is sketched in section 3, see in particular section 3.4, as well as appendix A. As will be explained below in section 3.5, these functions

#	Structure	Functions	$\Gamma_{g,h}$
1.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2A,2A_2}}$	$\Gamma(1)$
2.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2A,2A_3}}$	$\Gamma(1)$
3.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2A,2A_5}}$	$\Gamma(1)$
4.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2B,2B_2}}$	$\Gamma(1)$
5.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2B,2B_4}}$	$\Gamma(1)$
6.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2B,2B_6}}$	$\Gamma(1)$
7.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2B,2A_1}},\ \phi_{\mathrm{2A,2B_3}},\ \phi_{\mathrm{2A,2A_4}}$	$\Gamma_0(2)$
8.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2A,2B_1}},\ \phi_{\mathrm{2B,2A_2}},\ \phi_{\mathrm{2B,2B_1}}$	$\Gamma_0(2)$
9.	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\phi_{\mathrm{2A,2B_2}},\ \phi_{\mathrm{2B,2A_3}},\ \phi_{\mathrm{2B,2B_5}}$	$\Gamma_0(2)$
10.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4A_2}},\ \phi_{\mathrm{4A,2A_2}},\ \phi_{\mathrm{4A,4A_3}}$	$\Gamma_0(2)$
11.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4A_3}},\ \phi_{\mathrm{4A,2A_3}},\ \phi_{\mathrm{4A,4A_7}}$	$\Gamma_0(2)$
12.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4A_4}}, \phi_{\mathrm{2A,4B_4}}, \phi_{\mathrm{2B,4A_1}}, \phi_{\mathrm{2B,4B_1}}, \phi_{\mathrm{4A,2A_1}}, \phi_{\mathrm{4A,2B_1}},$ $\phi_{\mathrm{4B,2A_4}}, \phi_{\mathrm{4B,2B_2}}, \phi_{\mathrm{4B,4A_1}}, \phi_{\mathrm{4B,4A_2}}, \phi_{\mathrm{4A,4B_3}}, \phi_{\mathrm{4A,4B_4}}$	$\Gamma_{\mathrm{2A,4A}}$
13.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2B,4A_2}},\ \phi_{\mathrm{4A,2B_3}},\ \phi_{\mathrm{4A,4A_5}}$	$\Gamma_0(2)$
14.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2B,4A_3}},\ \phi_{\mathrm{4A,2B_2}},\ \phi_{\mathrm{4A,4A_2}}$	$\Gamma_0(2)$
15.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4C_1}}, \phi_{\mathrm{4C,2A_2}}, \phi_{\mathrm{4C,2B_3}}, \phi_{\mathrm{2B,4C_2}}, \phi_{\mathrm{4C,4C_4}}, \phi_{\mathrm{4C,4C_6}}$	$\Gamma_0(4)$
16.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4C_2}}, \phi_{\mathrm{4C,2A_1}}, \phi_{\mathrm{4C,2B_2}}, \phi_{\mathrm{2B,4C_3}}, \phi_{\mathrm{4C,4C_3}}, \phi_{\mathrm{4C,4C_5}}$	$\Gamma_0(4)$
17.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4B_2}},\ \phi_{\mathrm{4B,2A_2}},\ \phi_{\mathrm{4B,4B_8}}$	$\Gamma_0(2)$
18.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4B_3}},\ \phi_{\mathrm{4B,2A_5}},\ \phi_{\mathrm{4B,4B_9}}$	$\Gamma_0(2)$
19.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2A,4B_5}},\ \phi_{\mathrm{4B,2A_3}},\ \phi_{\mathrm{4B,4B_3}}$	$\Gamma_0(2)$
20.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2B,4B_2}},\ \phi_{\mathrm{4B,2B_3}},\ \phi_{\mathrm{4B,4B_1}}$	$\Gamma_0(2)$
21.	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\phi_{\mathrm{2B,4B_3}},\ \phi_{\mathrm{4B,2B_1}},\ \phi_{\mathrm{4B,4B_5}}$	$\Gamma_0(2)$
22.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\phi_{\mathrm{4A,4C_1}}, \phi_{\mathrm{4A,4C_2}}, \phi_{\mathrm{4C,4A_1}}, \phi_{\mathrm{4C,4A_2}}, \phi_{\mathrm{4C,4C_7}}, \phi_{\mathrm{4C,4C_8}}$	$\Gamma_{\mathrm{4A,4C}}$
23.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\phi_{\mathrm{4B,4A_3}},\ \phi_{\mathrm{4A,4B_1}},\ \phi_{\mathrm{4A,4A_1}}$	$\Gamma_0(2)$
24.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\phi_{\mathrm{4B,4A_4}},\ \phi_{\mathrm{4A,4B_2}},\ \phi_{\mathrm{4A,4A_4}}$	$\Gamma_0(2)$
25.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\phi_{\mathrm{4B,4B_4}}$	$\Gamma(1)$
26.	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\phi_{\mathrm{4B,4B_7}}$	$\Gamma(1)$
27.	$\mathbb{Z}_2 \times \mathbb{Z}_8$	$\phi_{\mathrm{2B,8A_{1,2}}}, \phi_{\mathrm{8A,2B_{1,2}}}, \phi_{\mathrm{4B,8A_{2,3}}}, \phi_{\mathrm{8A,4B_{1,3}}}, \phi_{\mathrm{8A,8A_{2,8}}}, \phi_{\mathrm{8A,8A_{6,7}}}$	$\Gamma_0(4)$
28.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$\phi_{\mathrm{2A,6A_2}}, \phi_{\mathrm{6A,2A_1}}, \phi_{\mathrm{6A,6A_1}}, \phi_{\mathrm{6A,6A_2}}$	$\Gamma_0(3)$
29.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$\phi_{\mathrm{2A,6A_3}}, \phi_{\mathrm{6A,2A_2}}, \phi_{\mathrm{6A,6A_3}}, \phi_{\mathrm{6A,6A_4}}$	$\Gamma_0(3)$
30.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$\phi_{\mathrm{2B,6B_2}}, \phi_{\mathrm{6B,2B_1}}, \phi_{\mathrm{6B,6B_2}}, \phi_{\mathrm{6B,6B_5}}$	$\Gamma_0(3)$
31.	$\mathbb{Z}_2 \times \mathbb{Z}_6$	$\phi_{\mathrm{2B,6B_3}}, \phi_{\mathrm{6B,2B_2}}, \phi_{\mathrm{6B,6B_1}}, \phi_{\mathrm{6B,6B_3}}$	$\Gamma_0(3)$
32.	$\mathbb{Z}_2 \times \mathbb{Z}_{10}$	$\phi_{\mathrm{2B,10A_1}}, \phi_{\mathrm{2B,10A_3}}, \phi_{\mathrm{10A,10A_1}}, \phi_{\mathrm{10A,10A_2}}, \phi_{\mathrm{10A,10A_3}}, \phi_{\mathrm{10A,10A_5}},$ $\phi_{\mathrm{10A,10A_7}}, \phi_{\mathrm{10A,10A_9}}, \phi_{\mathrm{10A,10A_{10}}}, \phi_{\mathrm{10A,10A_{11}}}, \phi_{\mathrm{10A,2B_2}}, \phi_{\mathrm{10A,2B_3}}$	$\Gamma_{\mathrm{2B,10A}}$
33.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\phi_{\mathrm{3A,3A_3}}$	$\Gamma(1)$
34.	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\phi_{\mathrm{3A,3B_1}}, \phi_{\mathrm{3B,3A_1}}, \phi_{\mathrm{3B,3B_3}}, \phi_{\mathrm{3B,3B_4}}$	$\Gamma_0(3)$

TABLE 2. The independent functions and the group $\Gamma_{g,h}$ for each $[g,h] \in \bar{P}$.

#	$\phi_{g,h}$	Obstr.
1.	0	1
2.	0	1
3.	0	1
4.	0	2
5.	0	2
6.	0	1
7.	0	1
8.	0	1
9.	0	1
10.	0	1
11.	0	1
12.	0	1
13.	$\phi_{2\mathrm{B},4\mathrm{A}_2} = 4\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	no
14.	0	1
15.	0	1
16.	0	1
17.	0	1
18.	0	1
19.	0	1
20.	0	2
21.	0	1
22.	0	2
23.	$\phi_{4\mathrm{B},4\mathrm{A}_3} = 2\sqrt{2}\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	no
24.	$\phi_{4\mathrm{B},4\mathrm{A}_4} = 2\sqrt{2}\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	no
25.	0	2
26.	0	2
27.	$\phi_{2\mathrm{B},8\mathrm{A}_{1,2}} = 2\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2$	no
28.	0	2
29.	0	2
30.	0	2
31.	0	2
32.	0	2
33.	$\phi_{3\mathrm{A},3\mathrm{A}_3} = 0$	no
34.	$\phi_{3\mathrm{A},3\mathrm{B}_1} = 0$	no

TABLE 3. For each of the 34 cases we give here the explicit result for the twisted twining genus, as well as where applicable, the obstruction that it responsible for its vanishing.

are also compatible with the expected projective representation of the centraliser $C_{M_{24}}(g)$ on the g-twisted sector, see property (C) in section 2.

3. HOLOMORPHIC ORBIFOLDS AND GROUP COHOMOLOGY

In this section we review the modular properties of holomorphic orbifolds since, according to our assumption (E), these are also relevant for the twisted twining genera of K3. We then explain how, in our context, the known multiplier phases of the twining genera determine the underlying 3-cocycle $\alpha \in H^3(M_{24}, U(1))$ uniquely, and how to obtain from this our explicit results given in table 3. Finally, we shall check that these formulae indeed give rise to the appropriate projective representations of the centraliser $C_{M_{24}}(g)$.

3.1. **Review of holomorphic orbifolds.** Suppose $\mathcal{C}$ is a self-dual VOA, and G is a group of $\mathcal{C}$-automorphisms. We are interested in the 'orbifold' of $\mathcal{C}$ by G. To this end we consider the G-invariant sub-VOA of $\mathcal{C}$,

$$\mathcal{C}^G = \{\psi \in \mathcal{C} : g\psi = \psi, \, \forall g \in G\} \ . \tag{31}$$

Each representation of $\mathcal{C}$ gives rise to a representation of $\mathcal{C}^G$. In addition, there are new representations of $\mathcal{C}^G$ that appear in the different twisted sectors $\mathcal{C}_A$; since $\mathcal{C}$ is holomorphic there is a unique twisted sector $\mathcal{C}_A$ for each conjugacy class A in G.

3.1.1. *Twisted twining characters.* Each twisted sector $\mathcal{C}_A$ defines a representation of the G-invariant sub-VOA $\mathcal{C}^G$, but it is typically not irreducible. In particular, for each g_A in the conjugacy class $g_A \in [A]$, we can project $\mathcal{C}_{g_A}$ onto a $\mathcal{C}^G$-invariant subspace using any (projective) character χ of $C_G(g_A)$. Thus we can label the irreducible representations of $\mathcal{C}^G$ by pairs (A, χ). In physics, by the orbifold of $\mathcal{C}$ by G one usually means the theory where one chooses $\chi = 1$ in each twisted sector, i.e. the full partition function takes the form

$$Z_{\rm orb}(\tau) = \sum_A \frac{1}{|C_G(A)|} \sum_{h \in C_G(g_A)} Z_{g_A,h}(\tau) \ , \tag{32}$$

where

$$Z_{g,h}(\tau) \equiv {}_g\underset{h}{\boxed{}} = \mathrm{Tr}_{\mathcal{C}_g} \left(\rho_g(h) \, q^{L_0 - \frac{c}{24}} \right) \ , \tag{33}$$

and ρ_g denotes the action of $C_G(g)$ on $\mathcal{C}_g$. We should mention that this action is usually *not* canonically defined; for example, what is usually referred to as discrete torsion [33] can be thought of as being an ambiguity in the definition of ρ_g. Incidentally, the fact that the action is not canonically defined will be important to us later, see in particular section 4.1. For the moment though, we want to assume that such an action has been chosen.

3.1.2. *Projective representations.* As alluded to above, the twisted sectors $\mathcal{C}_g$ typically do not form genuine representations of $C_G(g)$ but only *projective representations.* A projective representation ρ of a finite group G is characterised by the relation (see also appendix C for some introductory comments)

$$\rho(g_1) \, \rho(g_2) = c(g_1, g_2) \, \rho(g_1 g_2) \ , \tag{34}$$

where $c : G \times G \to U(1)$ is a (normalised) 2-cocycle representing a cohomology class in $H^2(G, U(1))$. Genuine representations of finite groups are classified, up to equivalence, by their characters $\mathrm{Tr}_\rho(h)$. This is, however, no longer true for projective representations since taking the trace over a projective representation does not lead to a class function, i.e. a function that is invariant under conjugation by elements in G. Instead, for projective representations one has

$$\mathrm{Tr}_\rho(hgh^{-1}) = \frac{c(h,h^{-1}gh)}{c(g,h)}\,\mathrm{Tr}_\rho(g)\ , \tag{35}$$

as follows from the defining relation (34) together with the cyclicity of the trace (see e.g. [34] for a nice discussion). For the case of the holomorphic orbifold, this then results in the conjugation relation

$$Z_{g,h}(\tau) = \frac{c_g(h,k)}{c_g(k,k^{-1}hk)}\,Z_{k^{-1}gk,k^{-1}hk}(\tau)\ , \tag{36}$$

where the 2-cocycle c_g represents a class in $H^2(C_G(g),U(1))$ specifying the projective representation of the centralizer $C_G(g)$ on $\mathcal{C}_g$, and we have assumed that k and g commute so that both expressions are evaluated in the same twisted sector. (In fact, (36) is even true if k and g do not commute, see [25].) If in addition k commutes with h, (36) reduces to

$$Z_{g,h}(\tau) = \frac{c_g(h,k)}{c_g(k,h)}\,Z_{g,h}(\tau)\ , \tag{37}$$

which therefore implies that $Z_{g,h}$ vanishes unless the 2-cocycle is *regular*,

$$c_g(h,k) = c_g(k,h)\ . \tag{38}$$

This is the holomorphic orbifold version of the first type of obstruction discussed in section 2.2 above.

3.1.3. *The underlying cohomology.* The possible choices for the 2-cocycles c_g in the various twisted sectors are constrained; for example this follows from demanding consistency of the OPE involving fields from different twist sectors.[‖] In particular, it was argued in [25] that the various consistent choices are in one-to-one correspondence with elements in the third cohomology group $[\alpha] \in H^3(G,U(1))$, see also [35, 36, 37, 38] for subsequent work. The 3-cocycle α determines distinguished elements $c_g \in H^2(C_G(g),U(1))$ via the formula [25, 35]

$$c_g(h_1,h_2) = \frac{\alpha(g,h_1,h_2)\,\alpha(h_1,h_2,(h_1h_2)^{-1}g(h_1h_2))}{\alpha(h_1,h_1^{-1}gh_1,h_2)}\ . \tag{39}$$

Since α fixes all 2-cocylces c_g via (39), it is clear that the choice of α determines the behaviour of the twisted twining characters under conjugation (36). Furthermore, α also

---

[‖] This origin of the constraint was suggested in [36].

determines all the multiplier phases that appear under modular transformations, since it follows from [35] that[**]

$$\begin{aligned} Z_{g,h}(\tau+1) &= c_g(g,h)\, Z_{g,gh}(\tau)\ , \\ Z_{g,h}(-1/\tau) &= \overline{c_h(g,g^{-1})}\, Z_{h,g^{-1}}(\tau)\ . \end{aligned} \tag{40}$$

The choice of a particular representative α in the class $[\alpha] \in H^3(G,U(1))$ corresponds to a choice of normalisation for the action of h on the g-twisted sector. To see this, let $\tilde{\alpha}$ represent a 3-cocycle in the same cohomology class $[\alpha] \in H^3(G,U(1))$, so that $\tilde{\alpha} = \alpha \cdot \partial\beta$ (see eq. (158)) for some 2-cochain β. Then the associate 2-cocycles c_g and $\tilde{c}_g$ differ (for each $g \in G$) only by a 2-coboundary i.e. $\tilde{c}_g = c_g \cdot \partial\gamma_g$, with $\gamma_g(h) = \frac{\beta(h,g)}{\beta(g,h)}$, see appendix B. Thus, the 3-cocycle $\tilde{\alpha}$ is associated with the modular properties of the functions $\tilde{Z}_{g,h} = \gamma_g(h) Z_{g,h}$, which, by (33), simply correspond to a different normalisation $\tilde{\rho}_g(h) = \gamma_g(h)\rho_g(h)$ for the action of h on the g-twisted sector.

3.2. **Application to Mathieu Moonshine.** Given what we said before, it is now natural to postulate that for the twisted twining genera of K3, the various multiplier phases also come from a 3-cocycle $\alpha \in H^3(M_{24},U(1))$. This is to say, we postulate that the projective representation appearing in (C) is the one determined by c_g, given by (39). Furthermore, $\chi_{g,h}\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in (4) is determined by (40), and $\xi_{g,h}(k)$ in (5) agrees with (37), where in both cases c_g is again the function determined by (39) for a fixed $\alpha \in H^3(M_{24},U(1))$. More explicitly, we therefore propose that

$$\begin{aligned} &\phi_{g,h}\left(-\tfrac{1}{\tau},\tfrac{z}{\tau}\right) = \overline{c_h(g,g^{-1})}\, e^{2\pi i \frac{z^2}{\tau}}\, \phi_{h,g^{-1}}(\tau,z)\ , \\ &\phi_{g,h}(\tau+1,z) = c_g(g,h)\, \phi_{g,gh}(\tau,z)\ , \end{aligned} \tag{41}$$

with $c_g(h_1,h_2)$ given by (39), and

$$\phi_{g,h}(\tau,z) = \frac{c_g(h,k)}{c_g(k,k^{-1}hk)}\, \phi_{k^{-1}gk,k^{-1}hk}(\tau,z)\ , \qquad k \in M_{24}\ . \tag{42}$$

Moreover, we postulate that

$$\phi_{g,h^{-1}}(\tau,z) = c_g(h,h^{-1})\, \phi^*_{g,h}(\tau,z)\ , \tag{43}$$

where $\phi^*_{g,h}(\tau,z)$ is defined in (12).

[**] Note that (40) is obtained from [35] upon working with the inverse 3-cohomology element. We also thank Terry Gannon for explaining these formulae to us. It is believed (and we have verified this in simple cases) that (40) follows from the transformation rules for the irreducible orbifold characters, see [37, eqs. (5.23) and (5.24)], but we have not proven this in complete generality.

Combining modular transformations with conjugation one obtains a whole set of relations among the twisted twining genera, where the multipliers are given by

$$\chi_{g,h}(\gamma_1\gamma_2) = \chi_{g,h}(\gamma_1)\,\chi_{(g,h)\gamma_1}(\gamma_2)\ , \qquad \gamma_1,\gamma_2 \in SL(2,\mathbb{Z})\ , \tag{44}$$

$$\chi_{g,h}\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \frac{1}{c_h(g,g^{-1})}\ , \tag{45}$$

$$\chi_{g,h}\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = c_g(g,h)\ , \tag{46}$$

$$\chi_{g,h}\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = \frac{1}{c_g(h,h^{-1})}\ , \tag{47}$$

$$\xi_{g,h}(k) = \frac{c_g(h,k)}{c_g(k,k^{-1}hk)}\ . \tag{48}$$

Since $SL(2,\mathbb{Z})$ is generated by $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ with the only relations being $S^4 = 1$ and $(ST)^3 = S^2$, it follows that eqs. (44) – (47) determine the multiplier $\chi_{g,h}(\gamma)$, $\gamma \in SL(2,\mathbb{Z})$, provided that the following consistency conditions are satisfied

$$\chi_{g,h}(S^4) = 1\ , \qquad \chi_{g,h}((ST)^3) = \chi_{g,h}(S^2)\ . \tag{49}$$

Furthermore, since the actions of $SL(2,\mathbb{Z})$ and M_{24} on the set $\mathcal{P}$ commute with one another, there are two more consistency conditions

$$\xi_{g,h}(k)\,\chi_{k^{-1}gk,k^{-1}hk}(\gamma) = \chi_{g,h}(\gamma)\,\xi_{(g,h)\gamma}(k) \qquad \xi_{g,h}(k)\,\xi_{k^{-1}gk,k^{-1}hk}(k') = \chi_{g,h}(kk')\ , \tag{50}$$

for all $\gamma \in SL(2,\mathbb{Z})$ and $k,k' \in M_{24}$. In fact, conditions (49) and (50) hold for a generic group G, and can be easily checked using standard identities for the cocycle c_g, in particular, eqs. (3.2.5)–(3.2.9) of [36] (see also [35]).

Note that the charge conjugation operator $C = S^2$ acts on the genera $\phi_{g,h}(\tau,z)$ by flipping the sign of the second argument z. In fact, writing $\tau = -\frac{1}{\tilde\tau}$ and $z = \frac{\tilde z}{\tilde \tau}$, we get from twice applying (41)

$$\begin{aligned}\phi_{g,h}(\tau,z) = \phi_{g,h}(-\tfrac{1}{\tilde\tau},\tfrac{\tilde z}{\tilde\tau}) &= \frac{e^{2\pi i\frac{\tilde z^2}{\tilde\tau}}}{c_h(g,g^{-1})}\,\phi_{h,g^{-1}}(\tilde\tau,\tilde z) \\ &= \frac{e^{-2\pi i\frac{z^2}{\tau}}}{c_h(g,g^{-1})}\,\phi_{h,g^{-1}}(-\tfrac{1}{\tau},-\tfrac{z}{\tau}) = \frac{1}{c_h(g,g^{-1})c_{g^{-1}}(h,h^{-1})}\,\phi_{g^{-1},h^{-1}}(\tau,-z)\ ,\end{aligned} \tag{51}$$

so that

$$\chi_{g,h}\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = \frac{1}{c_h(g,g^{-1})c_{g^{-1}}(h,h^{-1})}\ . \tag{52}$$

Together with (13), this then leads to

$$\phi_{g,h}(\tau,z) = \frac{1}{c_h(g,g^{-1})\,c_{g^{-1}}(h,h^{-1})}\,\phi_{g^{-1},h^{-1}}(\tau,z)\ , \tag{53}$$

thus reproducing (17). This identity played a key role in the derivation of obstruction 2 in section 2.2.

While our proposal about the multiplier phases is certainly natural, we do not have a direct proof for it. However, the fact that with this assumption we shall find consistent answers is in our opinion very convincing evidence in favour of this ansatz.

3.3. Determining the cohomology class.

In order to have explicit formulae for the multipliers, the next step consists of determining $[\alpha] \in H^3(M_{24}, U(1))$. In this subsection, we will prove (with some computer support) that there exists a unique cohomology class $[\alpha]$ that reproduces (10) for all the *untwisted* twining genera $\phi_{e,g}(\tau, z)$. Since this is an overdetermined problem, the existence of such class already represents a non-trivial consistency check for our proposal. In the following subsections, we will then explain how the remaining twisted twining characters can be computed, once the class $[\alpha]$ is known.

First we note that the normalisation condition for α, see eq. (159), implies that $c_e(g,h) = 1$ for all $g, h \in M_{24}$. Thus $\phi_{e,g}(\tau+1, z) = \phi_{e,g}(\tau, z)$, $\phi_{e,g^{-1}}(\tau, z) = \phi_{e,g}(\tau, z)$ and $\phi_{e,g}(\tau, z) = \phi_{e,k^{-1}gk}(\tau, z)$ for all $k \in M_{24}$, as expected. Furthermore, the representation ρ_e in the decomposition (6) is a genuine representation of $C_{M_{24}}(e) = M_{24}$.

Next we observe that eq. (41) implies

$$\phi_{g^a,g^b}(\tau + o(g^a), z) = \kappa(g^a, g^b)\, \phi_{g^a,g^b}(\tau, z) \ , \tag{54}$$

where

$$\kappa(g^a, g^b) := \prod_{k=1}^{o(g^a)} c_{g^a}(g^a, g^{b+ak}) = \prod_{k=1}^{o(g^a)} \alpha(g^a, g^{b+ak}, g^a) \ . \tag{55}$$

More precisely, the phases $\kappa(g^a, g^b)$ are the multipliers of $\phi_{e,g}$ under certain parabolic elements of the form

$$\gamma \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \gamma^{-1} \in \Gamma_0(o(g)) \ , \tag{56}$$

where $\gamma \in SL(2, \mathbb{Z})$ is such that $(e, g)\gamma = (g^a, g^b)$ and $n = o(g^a)$. By comparing such a multiplier with the one expected from (10), one obtains the following condition

$$\kappa(g^a, g^b) = e^{-\frac{2\pi i}{\ell(g^a)}} \ , \tag{57}$$

for all $a, b \in \mathbb{Z}$, where $\ell(g^a)$ is the length of the smallest cycle of g^a when considered as a permutation of 24 objects. It is easy to verify that $\kappa(g^a, g^b)$ only depends on the cohomology class of the 3-cocycle α. Thus we have to prove that there is a unique class $[\alpha] \in H^3(M_{24}, U(1))$ such that (57) is satisfied for all $g \in M_{24}$, $a, b \in \mathbb{Z}$.

In order to show that such an α exists, we first need two basic facts about group cohomology. Given two finite groups G and H and a homomorphism $\iota : H \to G$, we can define the induced map ι^* from $U(1)$-cochains α for G to $U(1)$-cochains $\iota^*(\alpha)$ for H

$$\iota^*(\alpha)(g_1, \ldots, g_n) = \alpha(\iota(g_1), \ldots, \iota(g_n)) \ , \tag{58}$$

which induces a well-defined map on cohomologies ([43], Chapter III.8). A particular case is when $H = G$ and the map $H \to G$ is conjugation $\iota_k(g) = k^{-1}gk$ by an element $k \in G$. In this case, the induced map on the cohomology ι_k^* acts as the identity, i.e. $[\iota^*(\alpha)] = [\alpha]$ (see [43], Chapter III Proposition 8.3). A second interesting case is when H is a subgroup of G with ι the natural inclusion; in this case, $\mathrm{Res} \equiv \iota^*$ is called a restriction map. In particular, if H is a Sylow p-subgroup $\mathrm{Syl}_p\, G$ of G,[††] then the image of the restriction map

$$\mathrm{Res}_p : H^n(G, U(1)) \to H^n(\mathrm{Syl}_p(G), U(1)) \tag{59}$$

is isomorphic to the p-part of $H^n(G, U(1))$ ([43], Chapter III Theorem 10.3). For $G = M_{24}$, using that $H^3(M_{24}, U(1)) \cong \mathbb{Z}_{12}$ [32], this implies

$$\mathrm{Res}_2(H^3(M_{24}, U(1))) \cong \mathbb{Z}_4 \ , \qquad \mathrm{Res}_3(H^3(M_{24}, U(1))) \cong \mathbb{Z}_3 \ , \tag{60}$$

while $\mathrm{Res}_p(H^3(M_{24}, U(1)))$ is trivial for all primes $p > 3$.

With these preparations we can now turn to the proof that a unique class $[\alpha] \in H^3(M_{24}, U(1))$ satisfying (57) exists. From the invariance of the cohomology classes under the conjugation map ι_k^*, it follows that the multiplier in (57) only depends on the conjugacy class of g. Next we note that in order to compute $\kappa(g^a, g^b)$, it is sufficient to consider the restriction of the cocycle α to the group $\langle g \rangle$ generated by g. For a cyclic group $\mathbb{Z}_N$, the third cohomology $H^3(\mathbb{Z}_N, U(1))$ is isomorphic to $\mathbb{Z}_N$, and a set of cocycle representatives for the cohomology classes are given by [44, 37]

$$\omega_q(g^a, g^b, g^c) = e^{2\pi i \frac{qa}{o(g)^2}([b]+[c]-[b+c])} \ , \qquad q = 0, \ldots, o(g) - 1 \ , \tag{61}$$

where $[\cdot] : \mathbb{Z} \to \{0, \ldots, o(g) - 1\}$ is the reduction modulo $o(g)$.

Let us assume, without loss of generality, that $1 \le a < o(g) - 1$ and $0 \le b < a$, and that a divides $o(g)$; in fact, if this was not true, we can rewrite both g^a and g^b as powers of another

---

[††] Recall that a p-group (p prime) is a finite group with order a power of p; a Sylow p-subgroup of a group G is a p-subgroup of G that is maximal, i.e. that is not properly contained in any other p-subgroup. For each p, all Sylow p-subgroups of G are isomorphic and conjugate to one another inside G.

$\tilde{g}$ in $\langle g\rangle$ for which this is then true. The multiplier corresponding to a 3-cocycle ω_q is

$$\begin{aligned}\kappa(g^a,g^b) &= \prod_{k=0}^{o(g^a)-1}\omega_q(g^a,g^{b+ak},g^a) = e^{2\pi i\frac{qa}{o(g)^2}\sum_{k=0}^{o(g^a)-1}([b+ak]+[a]-[b+a(k+1)])} \qquad (62)\\ &= e^{2\pi i\frac{qa}{o(g)^2}([b]+o(g^a)[a]-[b+ao(g^a)])} = e^{2\pi i\frac{qa}{o(g)}}\ ,\end{aligned}$$

where we have used that $o(g^a)=o(g)/a$ when a divides $o(g)$. Thus, $\kappa(g^a,g^b)\equiv\kappa(g^a)$ is independent of b, and $\kappa(g^a)=\kappa(g)^a$ whenever a divides $o(g)$. The latter property is consistent with (57) thanks to the identity

$$\frac{1}{\ell(g^a)}\equiv\frac{a}{\ell(g)}\mod 1\ , \qquad (63)$$

that holds for all elements $g\in M_{24}$ and for all divisors a of $o(g)$.[*]
It therefore follows that it is sufficient to prove (57) for $a=1$ and $b=0$, and for one representative g for each conjugacy class. Furthermore, thanks to $\kappa(g^a)=\kappa(g)^a$, we can restrict to classes whose order is a prime power. In other words, we only need to check (57) for a Sylow p-subgroup $\mathrm{Syl}_p(M_{24})$ and for each prime p.

Using the software `GAP` [41] with the package `HAP` [42] implemented, we have computed a basis of cocycle representatives for $H^3(\mathrm{Syl}_2(M_{24}),U(1))$ and $H^3(\mathrm{Syl}_3(M_{24}),U(1))$. Note that any 3-cocycle in the image of Res_p has the property that, for any $g\in\mathrm{Syl}_p(M_{24})$, the corresponding multiplier $\kappa(g)$ depends only on the conjugacy class of g in M_{24}. We have verified that there is a unique $\mathbb{Z}_4$ subgroup of $H^3(\mathrm{Syl}_2(M_{24}),U(1))$ and a unique $\mathbb{Z}_3$ subgroup of $H^3(\mathrm{Syl}_3(M_{24}),U(1))$ satisfying this property. Therefore these subgroups must correspond to the restrictions $\mathrm{Res}_p(H^3(M_{24},U(1)))$, for $p=2$ and $p=3$, respectively. Finally, for both $p=2$ and $p=3$, we have verified that there is a unique cohomology class $[\alpha_p]\in\mathrm{Res}_p(H^3(M_{24},U(1)))$ that satisfies (57) for all g in $\mathrm{Syl}_p(M_{24})$. Thus, the unique class $[\alpha]\in H^3(M_{24},U(1))$ that has the property that $\mathrm{Res}_p([\alpha])=[\alpha_p]$ for all primes p, satisfies (57) for all $g\in M_{24}$.

We have thus proven that there exists a unique class $[\alpha]\in H^3(M_{24},U(1))$ satisfying (57). It remains to show that this class exactly reproduces (10) for all $\phi_{e,g}$. First, we notice that the multiplier (10) of $\phi_{e,g}$ is trivial whenever the class of g restricts to a class in M_{23}. Since $H^3(M_{23},U(1))$ is trivial, this is consistent with any choice of the cocycle α. Thus, we only need to check (10) for the classes that do not restrict to M_{23}. The corresponding characters

---

[*] Note that this property is special for M_{24}, and is not true for a generic permutation.

$\phi_{e,g}$ are Jacobi forms under $\Gamma_0(o(g))$, where the possible orders $o(g)$ are $2, 3, 4, 6$ and 12. For $N = 2, 3, 4$, the group $\Gamma_0(N)$ is generated by

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} , \qquad C = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} , \qquad \begin{pmatrix} 1 & 0 \\ -N & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} T^N \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{64}$$

while $\Gamma_0(6)$ is generated by T, C, together with

$$\begin{pmatrix} 1 & 0 \\ -6 & 1 \end{pmatrix} , \qquad \begin{pmatrix} -5 & 3 \\ -12 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} T^3 \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} . \tag{65}$$

The correct modular properties (10) under T and C are automatic for any normalised cocycle α. The remaining generators are of the form (56) with $(e, g)\gamma = (g^a, g^b)$ and $n = o(g^a)$, and hence (10) follows again from (57).

The group $\Gamma_0(12)$ is generated by T, C and[†]

$$\alpha_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 12 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \qquad \alpha_2 = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} , \tag{66}$$

$$\alpha_3 = \begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -6 & 1 \end{pmatrix} , \qquad \alpha_4 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} . \tag{67}$$

The generators $\alpha_1, \ldots, \alpha_4$ are of the form (56) with $(e, g)\gamma = (g^a, g^b)$, but only for α_1, α_2 is the condition $n = o(g^a)$ satisfied, and thus only in these cases is condition (57) sufficient to establish (10). The generators α_3, α_4 satisfy $(e, g)\alpha_i = (e, g^7)$, $i = 3, 4$, so one needs to conjugate by an element $k \in M_{24}$ such that $k^{-1}g^7k = g$. In order to compute the multiplier under these generators, one needs to determine the restriction of $[\alpha] \in H^3(M_{24}, U(1))$ to the subgroup of M_{24} generated by g and k. Using `GAP`, we computed a basis for the third cohomology of the group $\langle g, k\rangle$ and looked for those classes that satisfy condition (57) for all elements of the group. It turns out that there is only one such class, and hence it must correspond to the restriction of $[\alpha] \in H^3(M_{24}, U(1))$ to the cohomology of $\langle g, k\rangle$. Then it is straightforward to check that this class reproduces indeed the multipliers (10) under α_3 and α_4; this is true for both the M_{24} conjugacy classes of order 12. This completes the proof. The `GAP`-files containing the computations described above are available online [45].

[†] It is straightforward to check that $T, C, \alpha_1, \ldots, \alpha_4$ actually generate $\Gamma_0(12)$ by simply comparing them to a minimal set of generators obtained by some standard method, such as the Farey symbol algorithm.

3.4. **Computation of the twisted twining characters: an example.** In the previous subsection, we proved that there is a unique cohomology class $[\alpha] \in H^3(M_{24}, U(1))$ that reproduces the multipliers (10) for all the *untwisted* twining genera $\phi_{e,g}$. In this section, we shall explain how the remaining twisted twining genera $\phi_{g,h}$ can be determined, leading to the results collected in section 2.5.

The first step is to compute the obstructions and the multipliers for a given twisted twining genus $\phi_{g,h}$. To this end we note that it is sufficient to consider the restriction of the cocycle α to the normaliser $N(g,h)$ of $\langle g, h\rangle$. If this normaliser is contained in some Sylow p-subgroup, one can simply use the cocycle α_p determined in the previous subsection. Otherwise, one can use `GAP` to compute a basis for the cohomology of $N(g,h)$ and check which cohomology classes reproduce the correct multipliers (and in particular (57)) for all genera of the form $\phi_{e,k}$, $k \in N(g,h)$. In all cases we considered, there was only one cohomology class with this property, which must therefore be the restriction of the class $[\alpha] \in H^3(M_{24}, U(1))$ to the cohomology of $N(g,h)$. Once the restriction of $[\alpha]$ to $N(g,h)$ is determined, one can check the existence of obstructions for all 34 groups of table 1. The results are collected in table 3 and the `GAP`-files containing these computations are available online [45].

For the six unobstructed twisted twining genera $\phi_{g,h}$, one can compute the precise multiplier system. It turns out that this information is sufficient to determine $\phi_{g,h}$ up to normalisation. The normalisation can be fixed (up to a phase) by requiring that a decomposition (6) exists. In the remainder of this subsection, we illustrate how one of these non-zero twisted twining genera, namely the one associated to $g \in 2\mathrm{B}$ and $h \in 8\mathrm{A}$, can be determined explicitly; the other five cases are completely analogous and are dealt with in appendix A. Let us work with the conventions that the three permutations

$$(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23)(24)\ ,$$
$$(3,17,10,7,9)(4,13,14,19,5)(8,18,11,12,23)(15,20,22,21,16)(1)(2)(6)(24)\ ,$$
$$(1,24)(2,23)(3,12)(4,16)(5,18)(6,10)(7,20)(8,14)(9,21)(11,17)(13,22)(15,19)$$

generate $M_{24} \subset S_{24}$. As representative for $g \in 2\mathrm{B}$ we can then take

$$g = (1,10)(2,14)(3,8)(4,5)(6,22)(7,20)(9,18)(11,23)(12,24)(13,19)(15,16)(17,21)\ , \quad (68)$$

while for $h \in 8\mathrm{A}$ we consider

$$h = (2,14)(3,9,8,18)(4,6,21,19,15,24,20,11)(5,22,17,13,16,12,7,23)(1)(10)\ . \quad (69)$$

The elements g, h generate the abelian group $\mathbb{Z}_2 \times \mathbb{Z}_8$ corresponding to case 27 of table 1. There are two interesting conjugation relations, namely

$$(k^{-1}gh^4k\,,\,k^{-1}h^{-1}k) = (g,h)\ , \qquad (r^{-1}gr\,,\,r^{-1}ghr) = (g,h)\ , \tag{70}$$

where

$$\begin{aligned} k &= (1,10)(2,14)(3,18)(4,11)(5,13)(6,20)(7,12)(8,9)(15,19)(16,23)(17,22)(21,24)\ ,\\ r &= (1,2)(3,9)(6,22)(8,18)(10,14)(11,23)(12,24)(13,19)(4)(5)(7)(15)(16)(17)(20)(21)\ . \end{aligned} \tag{71}$$

Thus, we can use the general formulae (41) and their inverses to deduce that[‡]

$$\phi_{g,h}(\tau+1) = c_g(g,h)\phi_{g,gh}(\tau) = c_g(g,h)\frac{c_g(gh,r)}{c_g(r,h)}\phi_{g,h}(\tau) = i\phi_{g,h}(\tau)\ , \tag{72}$$

where we have used the explicit form of the multiplier factors from section 3.3. We also have (recall that $g = g^{-1}$)

$$\begin{aligned} \phi_{g,h}(\tfrac{\tau}{-4\tau+1}) &= \frac{\phi_{h,g}(4-\frac{1}{\tau})}{c_h(g,g)} = \frac{\prod_{i=1}^{3} c_h(h,gh^i)}{c_h(g,g)}\phi_{h,gh^4}(-\tfrac{1}{\tau}) = \frac{\prod_{i=1}^{3} c_h(h,gh^i)}{c_h(g,g)c_{gh^4}(h,h^{-1})}\phi_{gh^4,h^{-1}}(\tau)\\ &= \frac{\prod_{i=1}^{3} c_h(h,gh^i)}{c_h(g,g)c_{gh^4}(h,h^{-1})}\frac{c_{g^4h}(h^{-1},k)}{c_{g^4h}(k,h)}\phi_{g,h}(\tau) = -\phi_{g,h}(\tau)\ . \end{aligned} \tag{73}$$

This implies that $\phi_{g,h}(\tau,z)$ is a Jacobi form under the group $\Gamma_0(4)$ with the above multipliers. Let us consider now the dependence of $\phi_{g,h}(\tau,z)$ on z, for a fixed $\tau \in \mathbb{H}_+$. The characters $\mathrm{ch}_{h,\ell}(\tau,z)$ of the $\mathcal{N}=4$ algebra are even functions of z and $\mathrm{ch}_{h,\ell}(\tau,z=0)$ is independent of τ (in particular, $\mathrm{ch}_{h,\ell}(\tau,0)=0$ for $h>1/4$) [31]. By (6), all twisted-twining genera must satisfy the same properties. Since the multiplier system for $\phi_{g,h}$ is non-trivial, the constant $\phi_{g,h}(\tau,0)$ must vanish and by $\phi_{g,h}(\tau,z)=\phi_{g,h}(\tau,-z)$ it follows that $\phi_{g,h}$ has a double zero at $z=0$ (mod $\mathbb{Z}+\tau\mathbb{Z}$). Every (weak) Jacobi form of index 1 has exactly two zeros (counting multiplicity) in each fundamental domain of $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ [30], so that these are the only zeroes of $\phi_{g,h}(\tau,z)$. On the other hand, the function

$$\psi(\tau,z) \equiv \frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2} \tag{74}$$

is also a Jacobi form weight 0 and index 1 under $\Gamma_0(4)$ (with trivial multiplier) and has only a double zero at $z=0$ (mod $\mathbb{Z}+\tau\mathbb{Z}$). It therefore follows that

$$F(\tau) \equiv \frac{\phi_{g,h}(\tau,z)}{\psi(\tau,z)} \tag{75}$$

[‡] For brevity of presentation we omit the dependence of z, as well as the factors $e^{2\pi i\frac{cz^2}{c\tau+d}}$, in the following formulae.

is independent of z and is a modular function of weight 0 under $\Gamma_0(4)$ (with multiplier system (72)–(73)) that is holomorphic in the interior of $\mathbb{H}_+$. Let us consider the asymptotic behaviour of $F(\tau)$ at the boundary $\Gamma_0(4)\backslash(\mathbb{Q}\cup\{\infty\})$ of the compactification $\Gamma_0(4)\backslash\bar{\mathbb{H}}_+$. This boundary consists of three points (cusps) of the form $\gamma_c\cdot\infty$, where $\gamma_\infty = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$, $\gamma_0 = \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ and $\gamma_{1/2} = \left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$. The twisted twining genus $\phi_{g,h}$ has the Fourier expansion at these cusps

$$\phi_{g,h}(\gamma_\infty(\tau,z)) = \phi_{g,h}(\tau,z) = A_\infty\, q^{1/4}(2-y-y^{-1}) + O(q^{5/4}) \tag{76}$$

$$\phi_{g,h}(\gamma_0(\tau,z)) \sim \phi_{h,g}(\tau,z) = A_0\, q^{1/8}(2-y-y^{-1}) + O(q^{3/8}) \tag{77}$$

$$\phi_{g,h}(\gamma_{1/2}(\tau,z)) \sim \phi_{gh^2,h}(\tau,z) = A_{1/2}\, q^{1/4}(2-y-y^{-1}) + O(q^{3/4})\ , \tag{78}$$

where $\sim$ denotes equality up to a phase, and we have included the lowest non-negative powers of q^r that are compatible with the multiplier system. It follows from (6) that the constants A_∞, A_0, $A_{1/2}$ equal

$$A_\infty = \mathrm{Tr}_{\mathcal{H}_{g,r=\frac{1}{4}}}(\rho_{g,\frac{1}{4}}(h))\ , \quad A_0 = \mathrm{Tr}_{\mathcal{H}_{h,r=\frac{1}{8}}}(\rho_{h,\frac{1}{8}}(g))\ , \quad A_{1/2} = \mathrm{Tr}_{\mathcal{H}_{gh^2,r=\frac{1}{4}}}(\rho_{gh^2,\frac{1}{4}}(h))\ , \tag{79}$$

and from the Fourier expansion of $\phi_{h,e}$ we know that $\dim\mathcal{H}_{h,r=\frac{1}{8}} = 1$, so that $|A_0| = 1$. On the other hand, at the cusps $\psi(\tau,z)$ satisfies

$$\psi(\gamma_\infty(\tau,z)) = \psi(\gamma_{1/2}(\tau,z)) = \frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2} = \frac{1}{4}(2-y-y^{-1}) + O(q) \tag{80}$$

$$\psi(\gamma_0(\tau,z)) = \frac{\vartheta_1(\tau,z)^2}{\vartheta_4(\tau,0)^2} = q^{1/4}(2-y-y^{-1}) + O(q^{3/4})\ , \tag{81}$$

and we therefore obtain

$$F(\gamma_\infty\cdot\tau) = 4\,A_\infty\, q^{1/4} + O(q^{5/4}) \tag{82}$$

$$F(\gamma_0\cdot\tau) \sim A_0\, q^{-1/8} + O(q^{1/8}) \tag{83}$$

$$F(\gamma_{1/2}\cdot\tau) \sim 4\,A_{1/2}\, q^{1/4} + O(q^{3/4})\ . \tag{84}$$

Up to a phase, there is a unique modular form with the correct modular properties and the expected Fourier expansion at the cusps, namely

$$F(\tau) = 8\frac{\eta(2\tau)^6}{\eta(\tau)^6}\ . \tag{85}$$

Thus we conclude that, up to an overall phase, we have

$$\phi_{2\mathrm{B},8\mathrm{A}_{1,2}}(\tau,z) = 8\frac{\eta(2\tau)^6}{\eta(\tau)^6}\frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2} = 2\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2\ , \tag{86}$$

where we used the identity $\vartheta_2(\tau,0)^2 = 4\frac{\eta(2\tau)^4}{\eta(\tau)^2}$. The other non-trivial cases can be worked out similarly; see appendix A for the details.

3.5. **Projective representations.** As we have argued above, see property (C) in section 2, the twisted sector $\mathcal{H}_g$ should carry a projective representation of the centraliser $C_{M_{24}}(g)$, whose 2-cocycle is given by c_g as determined in (39). Given our explicit knowledge of all twisted twining genera as well as c_g, this can now be tested as in [8] (where the corresponding analysis was performed for the case of the twining genera): if each $\mathcal{H}_{g,r}$ is a projective representation of $C_{M_{24}}(g)$, we can decompose it as

$$\mathcal{H}_{g,r} = \bigoplus_j h^{(j)}_{g,r}\, R_j \ , \tag{87}$$

where $h^{(j)}_{g,r}$ is the multiplicity with which the irreducible projective representation R_j (whose projectivity is characterised by c_g) appears in $\mathcal{H}_{g,r}$. On the other hand, using the orthogonality of group characters (see appendix C), we can calculate $h^{(j)}_{g,r}$ from the knowledge of the twisted twining genera explicitly. The consistency condition is then that all $h^{(j)}_{g,r}$ are indeed (non-negative) integers. We have done this analysis for all twisted sectors and for the first 500 levels (i.e. the first 500 values for r), and the multiplicities are indeed always non-negative integers. The explicit results for the first 20 levels as well as the (projective) character tables of all non-isomorphic centralisers $C_{M_{24}}(g)$ are given in appendix D. The decompositions for all g and for the first 500 levels are available online [45].

4. K3 Orbifolds

There are at least two further consistency checks on our proposal that can be fairly easily analysed. Suppose we consider a K3 sigma-model $\mathcal{C}$ whose automorphism group contains the cyclic subgroup generated by g. Then we can consider the orbifold of $\mathcal{C}$ by $G = \langle g\rangle$. A priori, it is not guaranteed that this orbifold is consistent — indeed, since these symmetries effectively act asymmetrically, they generically suffer from the level-matching problem, and hence may be inconsistent [39]. However, as was already mentioned in [40], the level-matching condition is satisfied provided that the twining genus ϕ_g has a trivial multiplier system. This is in particular the case if $g \in M_{23}$. In that case, the resulting orbifold theory $\widehat{\mathcal{C}} = \mathcal{C}/G$ is again a K3 sigma model, as was also shown in [40].

Suppose then that the original $\mathcal{C}$ has, in addition to $g \in M_{23}$, another commuting group element h in its automorphism group. Then h also gives rise to a symmetry $\hat{h}$ of the orbifold theory $\widehat{\mathcal{C}}$, and we can calculate the twining genus of $\hat{h}$ by the usual orbifold formula

$$\phi_{e,\hat{h}} = \frac{1}{o(g)} \sum_{i,j=0}^{o(g)-1} \phi_{g^i,hg^j} \ . \tag{88}$$

This leads to a non-trivial consistency condition: given our explicit knowledge of all twisted twining genera, we can calculate the right-hand-side explicitly, and this must agree, for every h, with *one* of the twining genera of [8, 9]. We have checked that this is indeed true; an example is illustrated in section 4.2.

Naively, one may have guessed that $\hat{h}$ should simply agree with h, i.e. that the left-hand-side of (88) is the twining genus associated to the *same* group element $h \in M_{24}$. However, this is in general not the case, as will be explained in the following section 4.1, see also section 4.2 for an explicit example. Whenever this happens we will say that the group element h has been *relabelled* in the orbifold theory.

The other consistency check is even more obvious. As was mentioned at the beginning of section 2, there exist K3 sigma-models with a commuting pair of automorphisms (g, h). For them (2) can be calculated directly, and thus compared to our answers. We shall perform this computation for some simple cases in section 4.3. In particular, we shall concentrate on examples where our obstruction analysis predicts that the associated twisted twining genus must vanish. The fact that we can reproduce this result using elementary methods is a good consistency check on our analysis.

4.1. **The relabelling phenomenon.**

Suppose that a K3 sigma-model $\mathcal{C}$ contains in its automorphism group an element g such that the orbifold of $\mathcal{C}$ by g, $\hat{\mathcal{C}} = \mathcal{C}/\langle g\rangle$ is consistent. In general, each g^r-twisted sector carries a projective representation of $H \equiv C_{M_{24}}(g)$. Equivalently, we can think of the untwisted and twisted sectors as carrying genuine representations of some central extension $\tilde{H}$ of H. In fact, the central extension can be chosen to be of the form (see also appendix C.1 for more details)

$$1 \to \langle\tilde{Q}\rangle \cong \mathbb{Z}_N \to \tilde{H} \xrightarrow{\rho} H \to 1 \ . \tag{89}$$

Here ρ is the representation in the untwisted sector, i.e. on the spectrum of $\mathcal{C}$, N is the order $o(g)$ of g and the generator $\tilde{Q}$ (that we call the *quantum symmetry*) of the central $\mathbb{Z}_N$ acts by $e^{2\pi i r/N}$ on the g^r-twisted sector. Note that there is always a lift $\tilde{g} \in \tilde{H}$ of g that acts by $e^{2\pi i k/N}$ on the states of conformal weight $h - 1/4 = k/N$.

In the orbifold theory $\hat{\mathcal{C}}$, all states are invariant under the action of $\tilde{g}$ — this is just the precise way of imposing the usual requirement that the orbifold theory consists of 'g-invariant' states only. Since $\tilde{g}$ is in the center of $\tilde{H}$, the spectrum of $\hat{\mathcal{C}}$ carries a well-defined representation $\hat{\rho}$ of $\tilde{H}$, such that $\hat{\rho}(\tilde{g})$ is the identity. In fact, it is easy to see that $\ker\hat{\rho} = \langle\tilde{g}\rangle$, because $\tilde{Q}$ acts non-trivially on the twisted sectors, and each non-trivial element in $H/\langle g\rangle$ lifts to an element of $\tilde{H}$ that acts non-trivially on the g-invariant untwisted sector, which in turn is part of the spectrum of $\hat{\mathcal{C}}$. Thus, we have an exact sequence

$$1 \to \langle\tilde{g}\rangle \cong \mathbb{Z}_N \to \tilde{H} \xrightarrow{\hat{\rho}} \hat{H} \to 1 \ . \tag{90}$$

The group $\hat{H} \equiv \tilde{H}/\langle\tilde{g}\rangle$ defines a full-fledged symmetry of the orbifold CFT $\hat{\mathcal{C}}$. For the case at hand, both the original theory and the orbifold are K3 sigma models, and thus all of these symmetry groups should be subgroups of M_{24},[§]

$$\hat{H} \subseteq M_{24} \ . \tag{91}$$

This assumption allows us to determine $\hat{H}$ (at least as an abstract group). In fact, the order of $\hat{H}$ is given by

$$|\hat{H}| = \frac{|\tilde{H}|}{N} = |H| = |C_{M_{24}}(g)| \ . \tag{92}$$

The quantum symmetry $Q \equiv \hat{\rho}(\tilde{Q})$ must be in the same M_{24}-conjugacy class as g, since it has the same eigenvalues on the 24-dimensional representation of RR ground states. Since all elements of $\hat{H}$ commute with Q, we can conclude by (92) that

$$\hat{H} = C_{M_{24}}(Q) \cong C_{M_{24}}(g) = H \ . \tag{93}$$

This construction allows us to give a precise interpretation of (88). For each symmetry $h \in H$ in the model $\mathcal{C}$, one chooses a lift $\tilde{h} \in \tilde{H}$, with $\rho(\tilde{h}) = h$ (this amounts to choosing the phases of the twisted twining genera $\phi_{g,h}$). Then the formula for the twining genus of the symmetry $\hat{\rho}(\tilde{h})$ in the orbifold model $\hat{\mathcal{C}}$ takes the form

$$\phi_{e,\hat{\rho}(\tilde{h})} = \frac{1}{o(g)} \sum_{i,j=0}^{o(g)-1} \phi_{g^i,\tilde{h}\tilde{g}^j} \ . \tag{94}$$

After these preparations we can now explain the relabelling phenomenon. While it follows from (93) that $\hat{H}$ and H are isomorphic, they are not canonically isomorphic; in fact, we only have a canonical isomorphism between

$$H/\langle g\rangle \cong \tilde{H}/\langle\tilde{g},\tilde{Q}\rangle \cong \hat{H}/\langle Q\rangle \ , \tag{95}$$

[§] We are assuming here that the relevant K3 sigma model is 'non-exceptional' in the terminology of [40].

but this does not necessarily lift to a canonical isomorphism $\Phi : \hat{H} \to H$, with $\Phi(Q) = g$. Put differently, in general it is not possible to find an isomorphism $\Phi : \hat{H} \to H$, with $\Phi(Q) = g$ such that the following diagram commutes:

$$\begin{tikzcd}
& \tilde{H} \arrow[dl, "\hat{\rho}"'] \arrow[dr, "\rho"] & \\
\hat{H} \arrow[rr, dashed, "\Phi"] \arrow[dr] & & H \arrow[dl] \\
& \tilde{H}/\langle \tilde{g}, \tilde{Q} \rangle &
\end{tikzcd}$$

For example, this will be the case if for any lift $\tilde{h}$ with $\rho(\tilde{h}) = h$, the corresponding symmetry $\hat{\rho}(\tilde{h})$ is not in the same M_{24} conjugacy class as h. In that case, the twining genus obtained by formula (94) does not agree with $\phi_{e,h}$. Whenever this is the case, we shall say that the conjugacy class of h has been *relabelled* in the orbifold theory.

4.2. **An example.** As an explicit example of the relabelling phenomenon let us consider the case where g is in the class 2A of M_{24} (so that $N = o(g) = 2$). Then the group $\tilde{H}$ is a central extension of H by the $\mathbb{Z}_2$ group generated by $\tilde{Q}$. We denote by $\tilde{g}$ the lift of g in $\tilde{H}$, such that $\tilde{g}$ and $\tilde{Q}$ have the following eigenvalues

$\tilde{Q}$	$\tilde{g}$	sector	model
$+1$	$+1$	untwisted	$\mathcal{C}, \hat{\mathcal{C}}$
$+1$	-1	untwisted	$\mathcal{C}$
-1	$+1$	twisted, $q^{\mathbb{Z}}$	$\hat{\mathcal{C}}$
-1	-1	twisted, $q^{\mathbb{Z}+\frac{1}{2}}$	$-$

(96)

As expected from the discussion in the previous section, the group $\tilde{H}/\tilde{Q}$ is isomorphic to the centraliser in M_{24} of an element of class 2A.

Now consider an element $h \in H$ in class 8A, with $h^4 = g$. The trace of h over the 24-dimensional representation of ground states in the original model $\mathcal{C}$ is $+2$. Furthermore, since hg is in the same class, also the trace over hg equals $+2$, from which we conclude that on the ground states of $\mathcal{C}$ we have

$$\mathrm{Tr}_{\mathcal{H}^0_{\tilde{Q}=1,\tilde{g}=1}}(h) = +2 \ , \qquad \mathrm{Tr}_{\mathcal{H}^0_{\tilde{Q}=1,\tilde{g}=-1}}(h) = 0 \ , \tag{97}$$

where $\mathcal{H}^0_{\tilde{Q}=\pm 1,\tilde{g}=\pm 1}$ are the eigenspaces of $\tilde{Q}$ and $\tilde{g}$ at conformal weight $h_L - \frac{1}{4} = 0$. The lifts $\tilde{h}$ and $\tilde{Q}\tilde{h}$ in $\tilde{H}$ are such that $\tilde{h}^4 = (\tilde{Q}\tilde{h})^4 = \tilde{g}$. On dimensional grounds, the twisted sector at $h_L - 1/4 = 0$ must correspond to a certain 8-dimensional irreducible representation of $\tilde{H}$, where the lifts $\tilde{h}$ and $\tilde{Q}\tilde{h}$ have trace ± 2. We notice that, up to a sign, this trace is independent of the choice of the central extension, and is in agreement with the expectation from the twisted twining genus $\phi_{g,h} \equiv \phi_{h^4,h}$. On the ground states in the orbifold theory $\hat{\mathcal{C}}$, the trace of $\tilde{h}$ is now

$$\mathrm{Tr}_{\mathcal{H}^0_{\tilde{Q}=1,\tilde{g}=1}}(\tilde{h}) + \mathrm{Tr}_{\mathcal{H}^0_{\tilde{Q}=-1,\tilde{g}=1}}(\tilde{h}) = 2 \pm 2 \;, \tag{98}$$

and thus there is no choice for $\tilde{h}$ such that the trace is equal to $+2$, as required for an element in class 8A. In fact, the images $\hat{h}$ and $Q\hat{h}$ in $\hat{H} \cong C_{M_{24}}(Q)$ have order 4 and belong to the classes 4A and 4B. In particular, while the pair of commuting elements g, h in the original model generate the cyclic group

$$\langle g, h\rangle = \langle h\rangle \cong \mathbb{Z}_8 \;, \tag{99}$$

both the groups $\langle Q, \hat{h}\rangle$ and $\langle Q, Q\hat{h}\rangle$ correspond to group 12 in our general list. Thus, there cannot exist an isomorphism $\Phi : \hat{H} \to H$ such that $\Phi(\hat{\rho}(\tilde{h})) = h$ or $\Phi(\hat{\rho}(\tilde{Q}\tilde{h})) = h$.

Let us verify the consistency of eq. (94) for this case, see the comments at the beginning of section 4. The twining genera $\phi_{e,\hat{\rho}(\tilde{h})}$ and $\phi_{e,\hat{\rho}(\tilde{Q}\tilde{h})}$ are given by

$$\phi_{e,\hat{\rho}(\tilde{h})}(\tau,z) = \frac{1}{2}(\phi_{e,h}(\tau,z) + \phi_{e,h^5}(\tau,z) + \phi_{h^4,h}(\tau,z) + \phi_{h^4,h^5}(\tau,z)) \;, \tag{100}$$

$$\phi_{e,\hat{\rho}(\tilde{Q}\tilde{h})}(\tau,z) = \frac{1}{2}(\phi_{e,h}(\tau,z) + \phi_{e,h^5}(\tau,z) - \phi_{h^4,h}(\tau,z) - \phi_{h^4,h^5}(\tau,z)) \;. \tag{101}$$

Using the identity

$$\phi_{e,8\mathrm{A}}(\tau,z) = \frac{1}{2}(\phi_{e,4\mathrm{B}}(\tau,z) + \phi_{e,4\mathrm{A}}(\tau,z)) \;, \tag{102}$$

it is easy to verify that

$$\phi_{e,\hat{\rho}(\tilde{h})} = \phi_{e,4\mathrm{B}}, \qquad \phi_{e,\hat{\rho}(\tilde{Q}\tilde{h})} = \phi_{e,4\mathrm{A}} \;, \tag{103}$$

as expected. This example also shows that the consistency of eq. (94) gives, in general, highly non-trivial conditions on the twisted twining genera.

We have verified explicitly that a similar phenomenon also occurs for the orbifold by an element g in class 4B.

4.3. **Computation of a twisted twining genus.** In this section we shall calculate some twisted twining genera directly. In particular, we shall concentrate on some simple cases where our general analysis predicts that they must vanish (because of some obstruction). As we shall see, we can reach the same conclusion from first principles.

Let us consider a K3 σ-model $\mathcal{C}$ whose automorphism group contains two commuting elements g, h of order m and n, respectively, generating the non-cyclic group $G = \langle g, h\rangle$. If these symmetries have a geometric origin, i.e. if they are induced by automorphisms of the K3 target space that fix the holomorphic 2-form, then the group G must be a subgroup of M_{23} with at least five orbits over the 24-dimensional representation, as follows from the Mukai theorem [46]. These conditions are satisfied by the groups 1 and 3 ($\mathbb{Z}_2 \times \mathbb{Z}_2$), groups 17 and 19 ($\mathbb{Z}_2 \times \mathbb{Z}_4$), group 25 ($\mathbb{Z}_4 \times \mathbb{Z}_4$), group 28 ($\mathbb{Z}_2 \times \mathbb{Z}_6$) and group 33 ($\mathbb{Z}_3 \times \mathbb{Z}_3$) of our table 1. In all these cases, the twisted twining genus vanishes, $\phi_{g,h} = 0$, see table 3. In the following we want to verify this independently.

In the general $\mathbb{Z}_n \times \mathbb{Z}_m$ case, the twisted twining genus $\phi_{g,h}$ is expected to be a weak Jacobi form of weight 0 and index 1 under the congruence subgroup

$$\Gamma_{m,n} \equiv \{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2,\mathbb{Z}) \mid a \equiv 1, b \equiv 0 \mod m,\ c \equiv 0, d \equiv 1 \mod n\}\ . \tag{104}$$

Note that this group is in general smaller than the group $\Gamma_{g,h}$ defined below (24), because we are not using the identifications under conjugation (5). Since both g and h are elements of M_{23}, the multiplier system for $\phi_{g,h}$ must be trivial (since $H^3(M_{23}, U(1)) \cong 0$).

Because of the triviality of the multiplier system, it follows that any modular image $\phi_{g',h'}$ of $\phi_{g,h}$, where $(g', h') = (g, h) \cdot \gamma$ for some $\gamma \in SL(2,\mathbb{Z})$, has a Fourier expansion of the form

$$\phi_{g',h'}(\tau, z) = \sum_{n=0}^{\infty} \sum_{\ell} c_{g',h'}(n, \ell)\, q^{\frac{n}{o(g')}}\, y^{\ell}\ , \tag{105}$$

where $o(g')$ is the order of g'. Let us assume

$$c_{g',h'}(0, \ell) = 0 \quad \forall \ell \in \mathbb{Z}\ , \tag{106}$$

for any modular image $(g', h') = (g, h) \cdot \gamma$; we will prove this below. Then it follows that

$$\phi_{g,h}(\tau, 0) = \sum_{\ell} c_{g,h}(0, \ell) = 0\ . \tag{107}$$

Because of (13), $\phi_{g,h}(\tau, z) = \phi_{g,h}(\tau, -z)$, and we can conclude that $\phi_{g,h}$ has a double zero at $z = 0$ (mod $\mathbb{Z} + \tau\mathbb{Z}$). Since Jacobi forms of index 1 have only two zeroes on each fundamental domain of $\mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ [30], $\phi_{g,h}$ is nowhere else vanishing. Next we recall that $\vartheta_1(\tau, z)^2/\eta(\tau)^2$ is a Jacobi form of weight 0 and index 1 under $SL(2,\mathbb{Z})$ (with a non-trivial multiplier). It has a double zero at $z = 0$ (mod $\mathbb{Z} + \tau\mathbb{Z}$), and is non-vanishing elsewhere, so that

$$F_{g,h}(\tau) = \frac{\eta(\tau)^2 \phi_{g,h}(\tau, z)}{\vartheta_1(\tau, z)^2} \tag{108}$$

is a modular function of weight 0 under $\Gamma_{m,n}$ (with non-trivial multiplier), which is holomorphic in the interior of the upper-half plane $\mathbb{H}_+$. Let us consider the asymptotics of $F_{g,h}$ at the boundary $\Gamma_{m,n}\backslash(\mathbb{Q} \cup \infty)$ of the compactification $\Gamma_{m,n}\backslash\bar{\mathbb{H}}_+$. The boundary is the union of points (cusps) of the form $\gamma(\infty)$, where γ runs through the representatives of the double cosets $\Gamma_{m,n}\backslash SL(2,\mathbb{Z})/\Gamma^\infty$. Here, Γ^∞ is the parabolic subgroup of $SL(2,\mathbb{Z})$ that fixes $\infty \in \bar{\mathbb{H}}_+$ and is generated by $\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$. The Jacobi form $\vartheta_1(\tau, z)^2/\eta(\tau)^2$ has the same asymptotic behaviour at each cusp, namely

$$\frac{\vartheta_1(\gamma \cdot (\tau, z))^2}{\eta(\gamma \cdot \tau)^2} \overset{\tau \to \infty}{\sim} q^{1/6}(2 - y - y^{-1}) + O(q^{7/6}) \ , \tag{109}$$

up to a constant coefficient. By (105) and (106), it then follows that

$$F_{g,h}(\gamma \cdot \tau) \sim \frac{\eta(\tau)^2 \phi_{g',h'}(\tau, z)}{\vartheta_1(\tau, z)^2} \overset{\tau \to \infty}{\sim} O(q^{\frac{1}{o(g')} - \frac{1}{6}}) \ , \tag{110}$$

where $(g', h') = (g, h)\gamma$ and $o(g')$ is the order of g'. For all the groups we are interested in, we have $o(g') \leq 6$, and the inequality is strict for at least one g'. It follows that $F_{g,h}$ has no poles and at least one zero at the boundary of $\Gamma_{m,n}\backslash\bar{\mathbb{H}}_+$. Since it is holomorphic in the interior, it must therefore vanish identically, and we conclude that

$$\phi_{g,h}(\tau, z) = 0 \ , \tag{111}$$

as predicted by our results.

It therefore only remains to prove (106). The coefficient $c_{g,h}(0, \ell)$ corresponds to the trace of h (or, more generally, of the lift $\tilde{h}$ to a central extension of the group $\langle g, h\rangle$) on the RR g-twisted states with conformal weight $h_L = h_R = 1/4$ and weight ℓ under the $\mathfrak{su}(2)_L$ subalgebra of the left-moving $\mathcal{N} = 4$ algebra. The g-twisted ground states are g-invariant (the conformal weight satisfies $h_L - 1/4 \in \mathbb{Z}$), so they belong to the spectrum of the orbifold theory $\mathcal{C}/\langle g\rangle$.

Any sigma-model whose target space is a K3 manifold X contains 24 RR ground states corresponding to the harmonic forms on X. Thus, the real vector space spanned by these states can be identified with the real cohomology $H^*(X, \mathbb{R})$ on X. The orbifold model $\hat{\mathcal{C}} = \mathcal{C}/\langle g\rangle$ corresponds to a non-linear sigma-model on the geometric orbifold $\hat{X} = X/\langle g\rangle$, which is a singular K3 surface. In particular, for each point of X that is fixed by some non-trivial subgroup $\langle g^r\rangle \subseteq \langle g\rangle$, the orbifold $\hat{X}$ has a singularity of type A_{n-1}, where n

is the order of g^r. By blowing up all these singularities, one obtains again a smooth K3 manifold.

The RR ground states in the untwisted sector of $\hat{\mathcal{C}}$ are obtained by projecting onto the g-invariant subspace of $H^*(X,\mathbb{R})$, while the ground states in the twisted sectors are the Poincaré duals of the cycles corresponding to the exceptional divisors in the blow-up of $\hat{X}$. More precisely, the resolution of each A_{n-1} singularity of $\hat{\mathcal{C}}$ gives rise to $n-1$ rational curves $\mathbb{P}^1$, representing elements in the homology of the blow-up. The cohomology classes dual to these cycles span an $(n-1)$-dimensional subspace of RR ground states for $\hat{\mathcal{C}}$, one in each g^{ir} twisted sector, for $i = 1, \ldots, n-1$. Thus, we arrive at the usual statement that the ground states of the g^k-twisted sector are in one to one correspondence with the g^k-fixed points of the target space.

Any other symplectic automorphism h of X that commutes with g acts as a permutation on the g^r-fixed points. This action induces a permutation on the singularities of $\hat{X}$, and therefore on the twisted RR ground states dual to the exceptional cycles.[¶] Therefore, the trace of h over the g-twisted sector only gets a non-zero contribution from the singularities of $\hat{X}$ that are fixed by h, i.e. only from the points of X that are fixed by both g and h. It is known (see, for example Proposition 1.5 of [46]) that a finite group of symplectic automorphisms of a K3 manifold X that fixes at least one point must be a subgroup of $SL(2,\mathbb{C})$. On the other hand, all finite abelian subgroups of $SL(2,\mathbb{C})$ are cyclic. Thus we conclude that if $\langle g, h\rangle$ is abelian but not cyclic, the trace of h over the g-twisted ground fields is 0 and (106) follows.

Our results also agree with the analysis of [47], where it was shown that the twisted twining genus $\phi_{g,h}$ for a pair of commuting symplectic automorphisms g, h of a K3 manifold X is given by a sum of contributions corresponding to the points fixed by both g and h. If the abelian group $\langle g, h\rangle$ is not cyclic, then there are no fixed points and $\phi_{g,h} = 0$.

5. CONCLUSIONS

5.1. **Summary.** Inspired by the generalised Monstrous Moonshine idea of Norton [14] we have in this paper established 'generalised Mathieu Moonshine' for the elliptic genus of K3. More specifically, we have found, for each commuting pair $(g,h) \in M_{24}$, explicit expressions for the twisted twining genera $\phi_{g,h} : \mathbb{H}_+ \times \mathbb{C} \to \mathbb{C}$, and we have verified that these functions satisfy conditions (A), (B), (D) of section 2.1, and given convincing evidence that condition (C) is also met. In particular, the twisted twining genera $\phi_{g,h}(\tau, z)$ are weak Jacobi forms of weight 0 and index 1 for certain subgroups $\Gamma_{g,h} \subset SL(2, \mathbb{Z})$ with a multiplier system $\chi_{g,h} : \Gamma_{g,h} \to U(1)$.

[¶] More precisely, one should choose a lift $\tilde{h}$ of h to a central extension of the group of symmetries of $\mathcal{C}$; then, $\tilde{h}$ acts on the g-twisted sector by a permutation times an overall phase that depends on the particular lift. Since the phase is not important for our argument, we will neglect this subtlety.

One of the key insights of our work is that these multiplier phases are all determined in terms of a class $[\alpha] \in H^3(M_{24}, U(1))$ via a formula (see eq. (39)) that was first derived by Dijkgraaf and Witten in the context of holomorphic orbifolds [25]. We have also shown that the twisted twining genera $\phi_{g,h}$ are compatible with a decomposition of the g-twisted sector into (projective) representations of the centraliser $C_{M_{24}}(g)$, up to the first 500 levels. The particular central extension of $C_{M_{24}}(g)$ associated with this projective representation is also determined by the cohomology class $[\alpha] \in H^3(M_{24}, U(1))$.

As it turned out, many of the twisted twining genera vanish, and in most cases this follows from the structure of the class $[\alpha] \in H^3(M_{24}, U(1))$. While these sorts of cohomological obstructions should also arise in the context of other holomorphic orbifolds (in particular in Monstrous Moonshine), our results provide, to our knowledege, the first example where they have been explicitly verified. We have also confirmed the vanishing of some of these twisted twining genera independently, using geometrical arguments.

5.2. **Open problems and future work.** The elliptic genus of K3 is closely related to the physics of 1/4 BPS-dyons in $\mathcal{N} = 4$ string theory. The generating function of elliptic genera of symmetric products of K3's is the famous Igusa cusp form Φ_{10}, whose inverse is the partition function of $\mathcal{N} = 4$ dyons [48, 49]. This fact lies at the heart of the recent progress in understanding wall-crossing of multi-centered BPS-states in $\mathcal{N} = 4$ string theory using mock modular forms [50]. It is therefore natural to wonder about the corresponding physical interpretation of the twisted twining genera $\phi_{g,h}$. In the special case of a trivial twist $g = e$, it has been shown that for some elements h the twining genera $\phi_{e,h}$ correspond to BPS-indices in CHL-orbifolds [6, 51, 52, 53]. We may therefore expect that the twisted twining genera $\phi_{g,h}$ should be related to the counting of 'twisted dyons' in CHL-models. Indeed for some pairs (g, h) we have verified this by showing that the twisted CHL-indices of Sen [47] are compatible with our $\phi_{g,h}$. It would be interesting to determine more generally whether $\phi_{g,h}$ have a physical interpretation in terms of CHL-models or some generalisation thereof.

In Borcherds' proof of the Monstrous Moonshine conjecture he introduced a new class of infinite-dimensional (super-)Lie algebras that he called generalised Kac-Moody algebras

(GKM) [3]. The key idea was that the multiplicative lift of the modular J-function could be interpreted as the denominator formula for a specific GKM, now commonly called the Monster Lie algebra, whose root system carries an action of the Monster group $\mathbb{M}$. The generalised moonshine ideas of Norton [14] suggest that similar GKMs should exist for each class $[g] \in \mathbb{M}$ [20, 21, 54]. In the context of Mathieu Moonshine the role of generalised Kac-Moody algebras has so far not been understood. Although there is a GKM associated with the elliptic genus of K3 [55], this algebra does not seem to carry any natural action of M_{24} due to the fact that its denominator formula is $1/\sqrt{\Phi_{10}}$ rather than $1/\Phi_{10}$. For some conjugacy classes $[g] \in M_{24}$ the associated second-quantised twining genera Φ_g also give rise to denominator formulas of GKMs [56], which should presumably be identified with the wall-crossing algebras found in CHL-models [57, 58, 59, 60, 51]. Although these observations are very suggestive, the precise role of the GKMs for Mathieu Moonshine remains to be understood.

A crucial point in our analysis was that the multiplier phases are controlled by a class $[\alpha] \in H^3(M_{24}, U(1))$, from which also the various obstructions could be read off. It follows from the work of [25, 37] that similarly a cohomology class in $H^3(\mathbb{M}, U(1))$ should underly generalised Monstrous Moonshine, and it would be very interesting to see whether this point of view may lead to new insights. In particular, some of Norton's generalised Moonshine functions $f(g, h; \tau)$ are known to vanish, and it would be natural to expect that this is due to some cohomological obstructions similar to those we have found in this paper.[‖] Some related ideas were already put forward some time ago by Mason [27], but very little has been done explicitly, mainly because the group $H^3(\mathbb{M}, U(1))$ seems to be poorly understood.

In a similar vein, it would also be interesting to investigate the generalised versions of the recently discovered Umbral Moonshine observations [61], of which the M_{24}/K3-moonshine is a special case. In particular, one might expect that there is a relation between the Rademacher summability condition of [12] and the cohomological obstructions we have described here.

The fact that there is a class in $H^3(M_{24}, U(1))$ determining the properties of the twisted twining genera $\phi_{g,h}$ represents strong evidence for the idea that a holomorphic VOA $\mathcal{H}$ underlies Mathieu Moonshine. Indeed, the original motivation for suspecting the relevance of the cohomology group $H^3(M_{24}, U(1))$ was based on the formal analogy with holomorphic orbifolds [25, 37]. In a sense we are therefore in a similar situation now as for Monstrous

---

[‖] We thank Terry Gannon for suggestions along these lines, see also [10].

Moonshine before the discovery of the Monster VOA $V^\natural$ [2]. However, the story appears to be more subtle since the natural expectation that $\mathcal{H}$ would arise as a superconformal σ-model on K3 does not hold [13]. On the other hand, this is perhaps not too surprising, given the fact that the modular J-function corresponds to the *full* partition function of a CFT, while the elliptic genus of K3 only receives contribution from a *subset* of the physical states of the $\mathcal{N} = (4,4)$ superconformal theory on K3. Hence it is tempting to speculate that $\mathcal{H}$ should correspond to something like the algebra of BPS-states of string theory on K3 (in the sense of Harvey and Moore [62, 63]). In particular, one might therefore hope that $\mathcal{H}$ can be constructed using some kind of topological sigma model, e.g. the 'half-twisted' topological σ-model on K3 [64]. Mathematically, this sigma model (or rather its large radius limit) can be viewed as a bundle of VOAs known as the chiral de Rham complex, Ω^{ch}_X [65], and in [66] it was shown that the elliptic genus of K3 can be obtained by taking the graded trace over the corresponding cohomology $H^*(\Omega^{ch}_X)$. It is therefore tempting to speculate that one may be able to construct an action of M_{24} on the cohomology of Ω^{ch}_X, at least for some choice of K3 surface X.

We hope to return to these and related issues in future work.

Acknowledgements

We are indebted to Terry Gannon for the suggestion that the 3rd cohomology of M_{24} should control the modular properties of the twisted twining genera, as well as for many useful explanations and directions to the relevant literature. We are furthermore grateful to Mathieu Dutour and Graham Ellis for helpful correspondence on using the computer package `HAP` for calculating the cohomology of finite groups. We also thank Nils Carqueville, Miranda Cheng, Jeff Harvey, Stefan Hohenegger, Axel Kleinschmidt, Sameer Murthy, Daniel Roggenkamp, Ashoke Sen and Don Zagier for useful discussions, as well as the anonymous referee for many useful suggestions that considerably improved the exposition of the paper. D.P. would like to thank the Albert Einstein Institute in Golm for hospitality while part of this work was carried out. The work of M.R.G. is partially supported by the Swiss National Science Foundation, and the visit of Henrik Ronellenfitsch at ETH was supported by the Pauli Center.

Appendix A. Details on the unobstructed twisted twining genera

In this appendix, following the example of section 3.4, we will determine the twisted twining genera $\phi_{g,h}$ for the groups 13, 23, 24, 33 and 34 in Table 1. The genus $\phi_{\mathrm{2B},\mathrm{8A}_{1,2}}$ for group

27 was obtained in section 3.4, while for all the other groups $\phi_{g,h}$ vanishes due to some obstruction.

A.1. **The characters $\phi_{2\mathrm{B},4\mathrm{A}_2}$ (group 13) $\phi_{4\mathrm{B},4\mathrm{A}_3}$ (group 23) and $\phi_{4\mathrm{B},4\mathrm{A}_4}$ (group 24).** In this section we will consider the twisted twining genera $\phi_{g,h}$ for groups 13, 23 and 24; as we will see, it turns out that these three characters have exactly the same modular properties. For group 13 g is in class 2B and h in class 4A, and we can choose

$$\begin{aligned} g &= (1,10)(2,14)(3,8)(4,16)(5,15)(6,12)(7,21)(9,18)(11,13)(17,20)(19,23)(22,24) \ , \\ h &= (1,2)(3,18)(4,7,15,17)(5,20,16,21)(6,19,24,11)(8,9)(10,14)(12,23,22,13) \ . \end{aligned} \tag{112}$$

For the groups 27 and 28, g is in class 4B and h in class 4A and we can choose

$$g = (3,8)(4,21,15,20)(5,17,16,7)(6,19,24,11)(9,18)(12,23,22,13) \ , \tag{113}$$

for both groups (from now on, we drop the trivial cycles). For group 23 we can take

$$h = (1,10)(2,14)(3,9)(4,11,5,13)(6,17,12,21)(7,22,20,24)(8,18)(15,19,16,23) \ . \tag{114}$$

and for group 24

$$h = (1,2,10,14)(3,9,8,18)(4,5,15,16)(6,13)(7,21,17,20)(11,22)(12,19)(23,24) \ , \tag{115}$$

In all the three cases we have the relations

$$(k^{-1}gk, k^{-1}ghk) = (g,h) \ , \qquad (r^{-1}gh^{-2}r, r^{-1}hr) = (g,h) \ , \tag{116}$$

where, for group 13

$$k = (1,9,2,3)(6,13,22,19)(7,21)(8,10,18,14)(11,24,23,12)(17,20) \ , \tag{117}$$

$$r = (3,8)(4,11,5,13)(6,20,12,7)(9,18)(15,19,16,23)(17,24,21,22) \ , \tag{118}$$

for group 23

$$k = (1,2,10,14)(3,9,8,18)(5,16)(6,11,24,19)(7,17)(12,23,22,13) \ , \tag{119}$$

$$r = (1,2)(3,8)(6,12)(7,20)(9,18)(10,14)(17,21)(22,24) \ , \tag{120}$$

and for group 24

$$k = (3,9)(4,12,5,6)(7,11,20,13)(8,18)(15,22,16,24)(17,19,21,23) \ , \tag{121}$$

$$r = (1,3)(2,9)(4,20)(5,7)(8,10)(14,18)(15,21)(16,17) \ . \tag{122}$$

Using the explicit knowledge of the 3-cocycle α we can derive the modular properties of the twisted twining genus $\phi_{g,h}$, including the multiplier system. It turns out that in all these three cases $\phi_{g,h}$ is invariant under $\Gamma_0(2)$ with the same multiplier

$$\phi_{g,h}(\tau+1) = c_g(g,h)\frac{c_g(gh,k)}{c_g(k,h)}\phi_{g,h}(\tau) = i\phi_{g,h}(\tau) \ , \tag{123}$$

$$\begin{aligned}\phi_{g,h}(\frac{\tau}{2\tau+1}) &= \frac{1}{c_h(h,h^{-1}g)c_h(h,h^{-2}g)}\frac{c_h(h^{-2}g,r)}{c_h(r,g)}\phi_{g,h}(\tau) \\ &= -i\phi_{g,h}(\tau) \ . \end{aligned}\tag{124}$$

By arguments similar to the ones in section 4.3, $\phi_{g,h}(\tau,z)$ has only a double zero at $z=0$ (mod $\mathbb{Z}+\tau\mathbb{Z}$). This implies that $\phi_{g,h}$ can be expressed as

$$\phi_{g,h}(\tau,z) = \frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2}F(\tau) \tag{125}$$

for some $\Gamma_0(2)$-invariant function $F(\tau)$ that is holomorphic in $\mathbb{H}_+$. The group $\Gamma_0(2)$ has two cusps at ∞ and 0. To proceed we study the asymptotic behavior at infinity

$$\phi_{g,h}(\tau,z) = A_\infty\, q^{1/4}(2-y-y^{-1}) + \cdots \tag{126}$$

$$\frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,0)^2} = \frac{1}{4}\,(2-y-y^{-1}) + \cdots \ , \tag{127}$$

which implies that $F(\tau) \sim 4A_\infty\, q^{1/4} + \cdots$ has a zero at $\tau\to\infty$. The behavior at $\tau=0$ is

$$\phi_{g,h}(\gamma_0(\tau,z)) \sim \phi_{h,g^{-1}}(\tau,z) = A_0\, q^{1/8}\,(2-y-y^{-1}) + \cdots \tag{128}$$

$$\frac{\vartheta_1(\tau,z)^2}{\vartheta_4(\tau,0)^2} = q^{1/4}\,(2-y-y^{-1}) + \cdots \ , \tag{129}$$

implying that $F(\gamma_0\cdot\tau) \sim A_0\, q^{-1/8} + \cdots$ has a pole at 0. A function with the correct modular properties and behaviour at the cusps is

$$F(\tau) = 4A_\infty\,\frac{\eta(2\tau)^6}{\eta(\tau)^6} \ , \tag{130}$$

so that

$$\phi_{g,h}(\tau,z) = 4A_\infty\frac{\eta(2\tau)^6}{\eta(\tau)^6}\frac{\vartheta_1(\tau,z)^2}{\vartheta_2(\tau,z)^2} = A_\infty\frac{\eta(2\tau)^2}{\eta(\tau)^4}\vartheta_1(\tau,z)^2 \ . \tag{131}$$

The constants A_∞ for the three cases are determined (up to the phase) by ensuring a decomposition into projective characters of $C_{M_{24}}(g)$. In particular, $A_\infty = 4$ for group 13 and $A_\infty = 2\sqrt{2}$ for groups 23 and 24.

A.2. **The cases $\phi_{3\mathrm{A},3\mathrm{A}_3}$ (group 33) and $\phi_{3\mathrm{A},3\mathrm{B}_1}$ (group 34).** In this section we show that the unobstructed twisted twining genera $\phi_{3\mathrm{A},3\mathrm{A}_3}$ and $\phi_{3\mathrm{A},3\mathrm{B}_1}$ must vanish by modular arguments.

Let us consider first the group $\langle g,h\rangle = \langle 3\mathrm{A}, 3\mathrm{B}_1\rangle$ (type 34), generated by

$$g = (2,23,11)(3,20,10)(5,14,9)(6,16,13)(7,19,22)(12,24,18) \tag{132}$$

in class 3A and

$$h = (1,15,17)(2,19,16)(3,18,5)(4,21,8)(6,11,7)(9,10,24)(12,14,20)(13,23,22) \tag{133}$$

in class 3B. The following conjugation relation holds

$$(k^{-1}gk, k^{-1}ghk) = (g,h)\ , \tag{134}$$

where

$$k = (2,7,16)(3,20,10)(4,21,8)(5,9,14)(6,11,22)(13,23,19)\ . \tag{135}$$

The modular properties of the associated twisted twining genus are

$$\phi_{g,h}(\tau+1) = c_g(g,h)\frac{c_g(gh,k)}{c_g(k,h)}\phi_{g,h}(\tau) = e^{\frac{4\pi i}{3}}\phi_{g,h}(\tau)\ , \tag{136}$$

and

$$\phi_{g,h}\left(\frac{\tau}{-3\tau+1}\right) = \prod_{i=0}^{2} c_h(h, h^i g^{-1})\phi_{g,h}(\tau) = e^{\frac{4\pi i}{3}}\phi_{g,h}(\tau)\ , \tag{137}$$

where we have used the explicit form for the multiplier phases as determined in terms of the 3-cocycle $\alpha \in H^3(M_{24}, U(1))$. We conclude that, up to the multiplier, $\phi_{g,h}$ is invariant under $\Gamma_0(3)$. Furthermore, as in the previous subsection and in section 4.3, for each fixed $\tau \in \mathbb{H}_+$, $\phi_{g,h}(\tau,z)$ has only a double zero at $z=0$ (mod $\mathbb{Z}+\tau\mathbb{Z}$). Since $\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^2}$ is a weak Jacobi form of weight 0 and index 1 with non-trivial multiplier under $SL(2,\mathbb{Z})$ and only a double zero at $z=0$, we have

$$\phi_{g,h}(\tau,z) = \frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^2}F_{g,h}(\tau) \tag{138}$$

for some function $F_{g,h}(\tau)$ that is $\Gamma_0(3)$-invariant (up to a multiplier). The asymptotics for $\tau \to \infty$ is

$$\phi_{g,h}(\tau,z) \sim A_\infty\, q^{\frac{2}{3}} + \cdots\ , \tag{139}$$

for some (possibly vanishing) constant A_∞, while the asymptotic behavior at 0 is given by

$$\phi_{h,g}(\tau,z) \sim A_0\, q^{\frac{2}{9}} + \cdots\ . \tag{140}$$

Since both at ∞ and 0 we have

$$\frac{\vartheta_1(\tau,z)^2}{\eta(\tau)^2} \sim q^{1/6}(2-y-y^{-1})+\cdots \ , \tag{141}$$

we deduce that $F_{g,h}$ has a zero at both 0 and ∞ and no poles; hence it must vanish. We therefore conclude

$$\phi_{3\mathrm{A},3\mathrm{B}_1}(\tau,z)=0 \ , \tag{142}$$

as claimed.

The $(3\mathrm{A},3\mathrm{A}_3)$-case (group 33) is analogous. Here, the group is generated by

$$g=(2,23,11)(3,20,10)(5,14,9)(6,16,13)(7,19,22)(12,24,18) \tag{143}$$

and

$$h=(2,22,13)(4,21,8)(5,14,9)(6,23,7)(11,19,16)(12,18,24) \ , \tag{144}$$

and we have the relations

$$(k^{-1}gk,k^{-1}ghk)=(g,h) \ , \qquad (r^{-1}hr,r^{-1}g^{-1}r)=(g,h) \tag{145}$$

where

$$k=(1,17,15)(3,9,24)(5,18,20)(6,16,13)(7,22,19)(10,14,12) \ , \tag{146}$$

$$r=(3,4,10,21)(5,18)(6,16,19,7)(8,20)(9,24,14,12)(11,22,23,13) \ . \tag{147}$$

The modular properties of $\phi_{g,h}$ are

$$\phi_{g,h}(\tau+1)=c_g(g,h)\,\frac{c_g(gh,k)}{c_g(k,h)}\,\phi_{g,h}(\tau)=e^{\frac{4\pi i}{3}}\,\phi_{g,h}(\tau) \ , \tag{148}$$

$$\phi_{g,h}(-1/\tau)=\frac{1}{c_h(g,g^{-1})}\frac{c_h(g^{-1},r)}{c_h(r,h)}\phi_{g,h}(\tau)=\phi_{g,h}(\tau) \ . \tag{149}$$

Thus $\phi_{g,h}$ is invariant (up to a multiplier) under $SL(2,\mathbb{Z})$, and $\phi_{g,h}\sim A\,q^{2/3}$ at ∞, so that

$$F_{3\mathrm{A},3\mathrm{A}_3}(\tau)\equiv\phi_{g,h}(\tau,z)\frac{\eta(\tau)^2}{\vartheta_1(\tau,z)^2}\sim A\,q^{1/2}+\cdots \ . \tag{150}$$

Thus, $F_{3\mathrm{A},3\mathrm{A}_3}$ has a zero and no poles and therefore must vanish, thus leading to

$$\phi_{3\mathrm{A},3\mathrm{A}_3}(\tau,z)=0 \ . \tag{151}$$

Appendix B. Some group cohomology

In this section, we will review the basic definitions and properties of group cohomology, see for example [43] for an introduction to the subject.

In general, group cohomology is defined in terms of a finite group G and a G-module A, i.e. an abelian group with an action of G satisfying certain compatibility properties. For our purposes, it is sufficient to consider $A = U(1)$ with the trivial action of G. An n-cochain ψ (with $n \geq 0$) for G with values in $U(1)$ is simply a function from $G \times \ldots \times G$ (n times) to $U(1)$; a 0-cochain is just an element of $U(1)$. On the space $C^n(G, U(1))$ of n-cochains one can define a *coboundary* operator $\partial_n : C^n(G, U(1)) \to C^{n+1}(G, U(1))$ by

$$\begin{aligned}\partial_n\psi\,(g_1, \ldots, g_{n+1}) &= \psi(g_2, \ldots, g_{n+1}) \\ &\times \prod_{i=1}^{n} \psi(g_1, \ldots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \ldots, g_{n+1})^{(-1)^i}\, \psi(g_1, \ldots, g_n)^{(-1)^{n+1}} \ . \quad (152)\end{aligned}$$

The cochains in the kernel of ∂_n are called n-cocycles (or co-closed n-cochains), while the cochains in the image of ∂_{n-1} are called n-coboundaries. The coboundary operator satisfies $\partial_{n+1} \circ \partial_n = 1$ (the trivial element of $U(1)$), so that we can define the cohomology, as usual, as the quotient of the space of cocycles modulo coboundaries

$$H^n(G, U(1)) = \frac{\ker \partial_n}{\mathrm{Im}\partial_{n-1}} \ , \tag{153}$$

where $n > 0$, and we use the convention that $H^0(G, U(1)) = U(1)$, and that every 0-cochain is coclosed, ie. $\partial_0 = 1$.

In particular, for a 1-cochain γ, the condition $\partial_1\gamma = 1$ is simply $\gamma(g)\gamma(h) = \gamma(gh)$ for all $g, h \in G$, so that $H^1(G, U(1))$ is just the group of homomorphism from G to $U(1)$. A 2-cochain $\beta : G \times G \to U(1)$ is co-closed (and hence defines a cocycle) provided it satisfies

$$\beta(g_1, g_2 g_3)\, \beta(g_2, g_3) = \beta(g_1 g_2, g_3)\, \beta(g_1, g_2) \tag{154}$$

for $g_1, g_2, g_3 \in G$. The second cohomology $H^2(G, U(1))$ then consists of the co-closed 2-cochains, modulo the ambiguity

$$\beta(g_1, g_2) \to \beta(g_1, g_2) \frac{\gamma(g_1)\gamma(g_2)}{\gamma(g_1 g_2)} \ , \tag{155}$$

where γ is an arbitrary 1-cochain, i.e. an arbitrary function $\gamma : G \to U(1)$. The second cohomology of a group G classifies the inequivalent projective representations of G, see appendix C. A 3-cochain α,

$$\alpha : G \times G \times G \to U(1) \tag{156}$$

is closed provided it satisfies

$$\alpha(g_1,g_2,g_3)\,\alpha(g_1,g_2g_3,g_4)\,\alpha(g_2,g_3,g_4)=\alpha(g_1g_2,g_3,g_4)\,\alpha(g_1,g_2,g_3g_4)\ . \tag{157}$$

In the cohomology group $H^3(G,U(1))$ the 3-cocycles are then identified modulo

$$\alpha(g_1,g_2,g_3)\to\alpha(g_1,g_2,g_3)\,\frac{\beta(g_1,g_2g_3)\beta(g_2,g_3)}{\beta(g_1g_2,g_3)\beta(g_1,g_2)}\ . \tag{158}$$

Note that the multiplying factor is trivial if β is closed, i.e. if it satisfies the 2-cocycle condition (154). We shall always work with normalised cocycles, i.e. we shall assume that

$$\alpha(e,g_1,g_2)=\alpha(g_1,e,g_2)=\alpha(g_1,g_2,e)=1 \tag{159}$$

for all $g_1,g_2\in G$, where e denotes the identity element of G.

Given a 3-cocycle α, we can define, for any $h\in G$, a 2-cochain $c_h:G\times G\to U(1)$ via

$$c_h(g_1,g_2)=\frac{\alpha(h,g_1,g_2)\,\alpha(g_1,g_2,(g_1g_2)^{-1}h(g_1g_2))}{\alpha(g_1,g_1^{-1}hg_1,g_2)}\ . \tag{160}$$

It is shown in [25] that c_h defines a 2-cocycle of the centraliser $C_G(h)\subset G$ (i.e. the subgroup of all elements g_1,g_2 which commute with h). When $g_1,g_2\in C_G(h)$, we have the simplified expression

$$c_h(g_1,g_2)=\frac{\alpha(h,g_1,g_2)\,\alpha(g_1,g_2,h)}{\alpha(g_1,h,g_2)}\ ,\qquad g_1,g_2\in C_G(h)\ . \tag{161}$$

Under the 'gauge transformation' (158), c_h transforms as

$$c_h(g_1,g_2)\to c_h(g_1,g_2)\frac{\gamma_h(g_1)\gamma_h(g_2)}{\gamma_h(g_1g_2)}\ ,\qquad g_1,g_2\in C_G(h)\ , \tag{162}$$

where we defined the 1-cochain γ_h by

$$\gamma_h(g)\equiv\frac{\beta(g,h)}{\beta(h,g)}\ . \tag{163}$$

This is indeed of the form (155), and hence, for all $h\in G$, c_h defines a map

$$c_h\ :\ H^3(G,U(1))\ \to H^2(C_G(h),U(1))\ . \tag{164}$$

Appendix C. Projective representations of finite groups

This section follows [37, 67]. A projective representation is a map $\rho:G\to\mathrm{End}(V)$ from a group G to the endomorphism group of some complex vector space V satisfying

$$\rho(g)\,\rho(h)=\beta(g,h)\,\rho(gh)\ ,\quad \beta(g,h)\in U(1)\ . \tag{165}$$

It follows from the associativity of the group multiplication that the phase $\beta(g,h)$ satisfies the 2-cocycle condition

$$\beta(g,h)\,\beta(gh,k)=\beta(h,k)\,\beta(g,hk)\ . \tag{166}$$

There is the freedom of rescaling the maps $\rho(g)$ as

$$\rho(g) \mapsto \gamma(g)\rho(g) \ , \quad \gamma(g) \in U(1) \ . \tag{167}$$

For the 2-cocylce β this leads to the rescaling by a coboundary

$$\beta(g,h) \to \beta(g,h)\,\gamma(gh)\,\gamma(g)^{-1}\,\gamma(h)^{-1} \ . \tag{168}$$

In particular, one can always choose a rescaling such that $\rho(e)$ is the identity matrix in $\mathrm{End}(V)$; with this choice, the cocycle β is normalised. Two projective representations whose 2-cocycles differ only by a coboundary are called projectively equivalent, since their representation matrices are related by a rescaling. Therefore, the inequivalent projective representations of G are captured by the second group cohomology $H^2(G,U(1))$.

As for genuine linear representations, one may define the character of a projective representation as the trace

$$\chi(g) = \ \mathrm{Tr}_V(\rho(g)) \ . \tag{169}$$

Due to the modified composition law, projective characters are in general not class functions, but differ by a phase on different representatives of one conjugacy class. However, it is always possible to choose a cocycle representative such that the characters are indeed class functions.

Assuming that $gh = hg$, we have, using the projective composition law,

$$\chi(g) = \chi(h^{-1}gh) = \frac{\beta(gh,h^{-1})}{\beta(h^{-1},gh)}\,\chi(g) \ . \tag{170}$$

Using the cocycle condition, it is easy to see that this condition on the characters is equivalent to the obstuction of the first kind (14). Also, the obstruction does not depend on the specific cocycle representative $\beta(g,h)$ or the class representative g. Classes whose projective characters vanish are called β-irregular. It may be shown that to each cocycle $\beta(g,h)$, there exist exactly as many linearly inequivalent irreducible projective representations as there are β-regular conjugacy classes.

Two irreducible characters belonging to representations with the same cocycle $\beta(g,h)$ satisfy an orthogonality relation of the form

$$\sum_{[g]} |K_g|^{-1}\,\chi(g)^*\,\tilde\chi(g) = \begin{cases} 1 & \chi = \tilde\chi \\ 0 & \text{otherwise,} \end{cases} \tag{171}$$

where the sum is taken over a set of (β-regular) class representatives of G, and $|K_g|$ is the size of the conjugacy class of g.

There is a natural connection between projective representations and central extensions of a group G. We say that $\tilde{G}$ is a central extention of G if there is an abelian subgroup $A \subset Z(\tilde{G})$ of the center $Z(\tilde{G})$ of $\tilde{G}$ such that the short sequence

$$0 \to A \to \tilde{G} \to G \to 0 \tag{172}$$

is exact, or equivalently that $G = \tilde{G}/A$. The simplest central extension is of course $\tilde{G} = A \times G$, but this is in general not the only possibility. It can be shown that the non-isomorphic central extensions of G by A are also classified by $H^2(G, A)$. Each set of projectively inequivalent projective representations of G corresponds to a set of genuine representations of some extension $\tilde{G}$.

In some sense there exists a maximal central extension S for each G, the Schur cover or 'Darstellungsgruppe'. It is characterised by the property that any projective representation of G corresponds to a genuine representation of S. The Schur cover is, in general, not unique, unless G is a perfect group, i.e. it coincides with the commutator subgroup. Otherwise, the Schur covers are related by isoclinism.

C.1. **Central extension and orbifolds.** Let $\mathcal{C}$ be a K3 sigma-model and suppose that its automorphism group contains an element g such that the orbifold of $\mathcal{C}$ by g, $\hat{\mathcal{C}} = \mathcal{C}/\langle g\rangle$ is consistent. In this subsection, we will show that all projective representations of $H \equiv C_{M_{24}}(g)$ on the g^r-twisted sectors are equivalent to genuine representations of a central extension $\tilde{H}$ of the form described in section 4.1, eq. (89).

By (54) and (62), the level-matching condition is satisfied if and only if the restriction of the class $[\alpha] \in H^3(M_{24}, U(1))$ to $H^3(\langle g\rangle, U(1))$ is trivial. Each g^r-twisted sector carries a, in general projective, representation ρ_{g^r} of $H \equiv C_{M_{24}}(g)$, with underlying 2-cocycle c_{g^r} given by (39). The various cocycles c_{g^r} are related to one another by the identity

$$c_{g^r}(h,k) = \frac{f_{g,r}(h) f_{g,r}(k)}{f_{g,r}(hk)}\, c_g(h,k)^r \ , \qquad h, k \in C(g) \ , \tag{173}$$

where

$$f_{g,r}(h) = \prod_{i=1}^{r-1} c_h(g, g^i) \ . \tag{174}$$

Here we have used the definition (39) of c_g and applied repeatedly the condition that the 3-cocycle α is closed. Thanks to the level-matching condition, we can choose the representative cocycle α to be identically 1 when restricted to $\langle g\rangle$, so that $f_{g,r}(g^i) = 1$ for all i. Furthermore, by changing α by a coboundary $\partial\beta$ as in (158), the functions $f_{g,r}(h)$ transform as

$$f_{g,r}(h) \to f_{g,r}(h)\, \frac{\beta(g^r,h)\beta(h,g)^r}{\beta(h,g^r)\beta(g,h)^r}\ . \tag{175}$$

Thus, we can choose β in such a way that $f_{g,r}(h) = 1$ for all $h \in C(g)$, so that

$$c_{g^r}(h,k) = c_g(h,k)^r\ , \qquad h,k \in C(g)\ . \tag{176}$$

In particular, since $c_e(h,k) = 1$, $c_g(h,k)$ is always an N-th root of unity, where $N = o(g)$.

APPENDIX D. CENTRALISERS $C_{M_{24}}(g)$ AND (PROJECTIVE) CHARACTER TABLES

This section describes the projective character tables of the centraliser $H \equiv C_{M_{24}}(g)$ for each conjugacy class of M_{24}, corresponding to the projective equivalence class determined by the 2-cocycle c_g. The procedure to compute these tables is as follows.

Let Schur(H) denote a Schur cover of H, i.e. a central extension of H fitting into the exact sequence

$$1 \to M(H) \to \mathrm{Schur}(H) \to H \to 1\ , \tag{177}$$

where $M(H) = H_2(H,\mathbb{Z}) = (H^2(H,U(1)))^*$ is the Schur multiplier. The Schur cover Schur(H) and its character table can be easily computed (we used `GAP` [41] for this). Any irreducible projective representation ρ of H is equivalent to a genuine irreducible representation $\tilde\rho$ of Schur(H) and vice versa. Thus, one can associate a class of $H^2(H,U(1))$ to each irreducible representation of Schur(H). The projective character table we are interested in is obtained from the one of Schur(H) by keeping only the representations associated to the class $[c_g]$, and by choosing a lift in Schur(H) of each conjugacy class of H (different choices of the lifts lead to projectively equivalent representations).

It only remains to determine which irreducible representations of Schur(H) correspond to the class $[c_g]$. Consider an irreducible projective representation ρ of H associated with the 2-cocycle c_g. For any commuting pair $h,k \in H$, $hk = kh$, we have

$$\rho(h)\rho(k)\rho(h)^{-1}\rho(k)^{-1} = \frac{c_g(h,k)}{c_g(k,h)}\ . \tag{178}$$

(Note that $\frac{c_g(h,k)}{c_g(k,h)}$ only depends on the class $[c_g]$ and not on the representative cocycle c_g.) Thus, any representation σ of Schur(H) in the class $[c_g]$ must satisfy

$$\sigma(\tilde h\tilde k\tilde h^{-1}\tilde k^{-1}) = \frac{c_g(h,k)}{c_g(k,h)}\ , \tag{179}$$

where $\tilde h, \tilde k$ are some lifts of h,k in Schur(H). It is easy to see that the commutator $\tilde h\tilde k\tilde h^{-1}\tilde k^{-1}$ is an element of $M(H)$, and that it does not depend of the choice of the lifts. In fact, the commutators of this form generate $M(H)$, and if a representation σ satisfies (179) for a set of generators of $M(H) = (H^2(H, U(1)))^*$, then it must be associated with the class $[c_g]$. This condition can be easily tested for each irreducible representation of Schur(H), given our explicit knowledge of the cocycle c_g.

D.1. **Character Tables.** In this section we collect some information about the centralisers $C_{M_{24}}(g)$ of representatives of each M_{24} conjugacy class $[g]$, in particular a (rough) description of the group structure and the group order. Our notation is as follows:

$G = N.Q$	group G containing a normal subgroup N and such that $G/N \cong Q$
$N \rtimes Q$	(left) semidirect product of N and Q
$\mathbb{Z}_n$	cyclic group of order n
S_n	group of permutations (symmetric group) of n objects
A_n	group of even permutations (alternating group) of n objects
D_n	dihedral group of order n
$PSL_n(q)$	projective special linear group over the n-dim. vector space over the finite field $\mathbb{F}_q$

Furthermore, for each class $[g]$, we specify whether the representation carried by the g-twisted sector $\mathcal{H}_g$ is (projectively equivalent to) a linear representation, or whether it is genuinely projective; in the latter case we describe a minimal central extension with respect to which the representation is linear. Finally, we provide the character table of $C_{M_{24}}$ for the relevant projective equivalence class. The names of the conjugacy classes are chosen to correspond to the ATLAS names of M_{24} classes, with additional indices in case of redundancy. We adopt the following notation for algebraic numbers

$$\varepsilon_n = e^{\frac{2\pi i}{n}} \ , \qquad\qquad e_p^{\pm} = \frac{1}{2}(1 \pm i\sqrt{p}) \ .$$

Classes 11A, 12A, 12B, 14AB, 15AB, 21AB, 23AB

When g is in one of these classes, the centralizer $C_{M_{24}}(g)$ is the cyclic group $\langle g \rangle$ and all representations in the g-twisted sector are (projectively equivalent to) linear representations. Thus, all irreducible representations are one dimensional, and a useful choice for the corresponding (projective) characters χ_k, $k = 1, \ldots, o(g)$ is

$$\chi_k(g^a) = e^{2\pi i \frac{a}{o(g)}(k - \frac{1}{\ell(g)})} \ , \qquad (180)$$

$a = 0, \dots, o(g) - 1$, where $\ell(g)$ is the length of the shortest cycle of g when considered as a permutation of 24 objects. Indeed, the characters (180) correspond to the standard definition $\rho_g(g^a) = e^{2\pi i a (L_0 - \tilde{L}_0)}$ for the action of $\langle g \rangle$ in the g-twisted sector.

Class 2A

$C_{M_{24}}(2A) \cong \mathbb{Z}_2^4.(\mathbb{Z}_2^3 \rtimes PSL_3(2))$ Order: $2^{10} \cdot 3 \cdot 7 = 21504$

Genuinely projective rep. Central extension: $\mathbb{Z}_2.C_{M_{24}}(2A)$

	21504	21504	1536	1536	384	512	256	256	64	128	128	64	128	64	32	64	32	32	16	16	16	24	24	12	12	12	14	14	14	14
	1A	$2A_1$	$2A_2$	$2A_3$	$4A_1$	$2B_1$	$2A_4$	$2B_3$	$4B_2$	$2A_5$	$2B_2$	$4A_2$	$4B_1$	$4A_3$	$4C_1$	$4B_3$	$4B_4$	$4A_4$	$4B_5$	$4C_2$	$8A_1$	$3A_1$	$6A_1$	$6A_2$	$6A_3$	$12A_1$	$7A_1$	$14A_1$	$7B_1$	$14B_1$
χ_1	8	-8	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	2	-2	2	0	0	0	-1	1	-1	1
χ_2	8	8	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-2	-2	0	0	-2	-1	-1	-1	-1
χ_3	8	8	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	$\sqrt{3}$	$\sqrt{3}$	1	-1	-1	-1	-1
χ_4	8	8	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	$-\sqrt{3}$	$-\sqrt{3}$	1	-1	-1	-1	-1
χ_5	24	-24	0	0	0	0	0	0	0	0	0	0	-4	0	0	0	0	0	0	0	2	0	0	0	0	0	e_7^-	$-e_7^-$	e_7^+	$-e_7^+$
χ_6	24	-24	0	0	0	0	0	0	0	0	0	0	-4	0	0	0	0	0	0	0	2	0	0	0	0	0	e_7^+	$-e_7^+$	e_7^-	$-e_7^-$
χ_7	24	24	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e_7^-	e_7^-	e_7^+	e_7^+
χ_8	24	24	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e_7^+	e_7^+	e_7^-	e_7^-
χ_9	48	-48	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	1	-1
χ_{10}	56	-56	0	0	0	0	0	0	0	0	0	0	-4	0	0	0	0	0	0	0	-2	-2	2	0	0	0	0	0	0	0
χ_{11}	56	56	0	0	-4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-2	-2	0	0	2	0	0	0	0
χ_{12}	56	56	0	0	-4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	$-\sqrt{3}$	$\sqrt{3}$	-1	0	0	0	0
χ_{13}	56	56	0	0	-4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	$\sqrt{3}$	$-\sqrt{3}$	-1	0	0	0	0
χ_{14}	64	-64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	-2	0	0	0	-1	1	-1	1

TABLE 4. Projective character table for the centralizer of class 2A.

Class 2B

$C_{M_{24}}(2B) \cong \mathbb{Z}_2^6 \rtimes S_5$ Order: $2^9 \cdot 3 \cdot 5 = 7680$

Genuinely projective rep. Central extension: $\mathbb{Z}_2.C_{M_{24}}(2B) \cong (\mathbb{Z}_2.\mathbb{Z}_2^6) \rtimes S_5$

	7680	512	256	3840	512	7680	96	32	96	32	24	24	12	12	12	128	64	32	32	64	64	128	20	20	20	20	16	16	16	16
	1A	2B$_1$	2A$_1$	2B$_2$	2A$_2$	2B$_3$	2B$_4$	4A$_1$	4C$_1$	4C$_2$	3B$_1$	6B$_1$	6B$_2$	6B$_3$	12B$_1$	2B$_5$	4A$_2$	4B$_1$	4A$_3$	4B$_2$	2B$_6$	2A$_3$	5A$_1$	10A$_1$	10A$_2$	10A$_3$	4C$_3$	8A$_1$	4B$_3$	8A$_2$
χ_1	4	0	0	4	0	−4	2	0	−2	0	−1	1	−1	−1	1	0	2	0	0	−2	0	0	−1	1	1	−1	0	$2e_1^-$	0	$2e_1^+$
χ_2	4	0	0	−4	0	−4	2	0	2	0	−1	1	1	−1	−1	0	2	0	0	2	0	0	−1	−1	1	1	0	$-2e_1^+$	0	$-2e_1^-$
χ_3	4	0	0	4	0	−4	−2	0	2	0	−1	1	−1	1	−1	0	2	0	0	−2	0	0	−1	1	1	−1	0	$-2e_1^-$	0	$-2e_1^+$
χ_4	4	0	0	−4	0	−4	−2	0	−2	0	−1	1	1	1	1	0	2	0	0	2	0	0	−1	−1	1	1	0	$2e_1^+$	0	$2e_1^-$
χ_5	16	0	0	16	0	−16	4	0	−4	0	−1	1	−1	1	−1	0	0	0	0	0	0	0	1	−1	−1	1	0	0	0	0
χ_6	16	0	0	−16	0	−16	4	0	4	0	−1	1	1	1	1	0	0	0	0	0	0	0	1	1	−1	−1	0	0	0	0
χ_7	16	0	0	16	0	−16	−4	0	4	0	−1	1	−1	−1	1	0	0	0	0	0	0	0	1	−1	−1	1	0	0	0	0
χ_8	16	0	0	−16	0	−16	−4	0	−4	0	−1	1	1	−1	−1	0	0	0	0	0	0	0	1	1	−1	−1	0	0	0	0
χ_9	16	0	0	0	0	16	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	1	$\sqrt{5}$	1	$\sqrt{5}$	0	0	0	0
χ_{10}	16	0	0	0	0	16	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	1	$-\sqrt{5}$	1	$-\sqrt{5}$	0	0	0	0
χ_{11}	20	0	0	20	0	−20	−2	0	2	0	1	−1	1	1	−1	0	2	0	0	−2	0	0	0	0	0	0	0	$2e_1^-$	0	$2e_1^+$
χ_{12}	20	0	0	−20	0	−20	−2	0	−2	0	1	−1	−1	1	1	0	2	0	0	2	0	0	0	0	0	0	0	$-2e_1^+$	0	$-2e_1^-$
χ_{13}	20	0	0	20	0	−20	2	0	−2	0	1	−1	1	−1	1	0	2	0	0	−2	0	0	0	0	0	0	0	$-2e_1^-$	0	$-2e_1^+$
χ_{14}	20	0	0	−20	0	−20	2	0	2	0	1	−1	−1	−1	−1	0	2	0	0	2	0	0	0	0	0	0	0	$2e_1^+$	0	$2e_1^-$
χ_{15}	24	0	0	24	0	−24	0	0	0	0	0	0	0	0	0	0	−4	0	0	4	0	0	−1	1	1	−1	0	0	0	0
χ_{16}	24	0	0	−24	0	−24	0	0	0	0	0	0	0	0	0	0	−4	0	0	−4	0	0	−1	−1	1	1	0	0	0	0
χ_{17}	32	0	0	0	0	32	0	0	0	0	−2	−2	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0
χ_{18}	48	0	0	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−2	0	−2	0	0	0	0	0

TABLE 5. Projective character table for the centralizer of class 2B.

Class 3A

$C_{M_{24}}(3A) \cong \mathbb{Z}_3.A_6$ Order: $2^3 \cdot 3^3 \cdot 5 = 1080$ Linear rep.

	1080	1080	1080	24	24	24	9	9	12	12	12	15	15	15	15	15	15
	1A	3A$_1$	3A$_2$	6A$_1$	6A$_2$	2A$_1$	3A$_3$	3B$_1$	12A$_1$	12A$_2$	4A$_1$	15A$_1$	5A$_1$	15B$_1$	5A$_2$	15A$_2$	15B$_2$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	3	$-3e_3^+$	$-3e_3^-$	e_3^+	e_3^-	-1	0	0	$-e_3^+$	$-e_3^-$	1	α	e_{-5}^+	$\bar{\alpha}$	e_{-5}^-	β	$\bar{\beta}$
χ_3	3	$-3e_3^+$	$-3e_3^-$	e_3^+	e_3^-	-1	0	0	$-e_3^+$	$-e_3^-$	1	$\bar{\beta}$	e_{-5}^-	β	e_{-5}^+	$\bar{\alpha}$	α
χ_4	3	$-3e_3^-$	$-3e_3^+$	e_3^-	e_3^+	-1	0	0	$-e_3^-$	$-e_3^+$	1	β	e_{-5}^-	$\bar{\beta}$	e_{-5}^+	α	$\bar{\alpha}$
χ_5	3	$-3e_3^-$	$-3e_3^+$	e_3^-	e_3^+	-1	0	0	$-e_3^-$	$-e_3^+$	1	$\bar{\alpha}$	e_{-5}^+	α	e_{-5}^-	$\bar{\beta}$	β
χ_6	5	5	5	1	1	1	-1	2	-1	-1	-1	0	0	0	0	0	0
χ_7	5	5	5	1	1	1	2	-1	-1	-1	-1	0	0	0	0	0	0
χ_8	6	$-6e_3^+$	$-6e_3^-$	$-2e_3^+$	$-2e_3^-$	2	0	0	0	0	0	$-e_3^-$	1	$-e_3^+$	1	$-e_3^+$	$-e_3^-$
χ_9	6	$-6e_3^-$	$-6e_3^+$	$-2e_3^-$	$-2e_3^+$	2	0	0	0	0	0	$-e_3^+$	1	$-e_3^-$	1	$-e_3^-$	$-e_3^+$
χ_{10}	8	8	8	0	0	0	-1	-1	0	0	0	e_{-5}^+	e_{-5}^+	e_{-5}^+	e_{-5}^-	e_{-5}^-	e_{-5}^-
χ_{11}	8	8	8	0	0	0	-1	-1	0	0	0	e_{-5}^-	e_{-5}^-	e_{-5}^-	e_{-5}^+	e_{-5}^+	e_{-5}^+
χ_{12}	9	9	9	1	1	1	0	0	1	1	1	-1	-1	-1	-1	-1	-1
χ_{13}	9	$-9e_3^+$	$-9e_3^-$	$-e_3^+$	$-e_3^-$	1	0	0	$-e_3^+$	$-e_3^-$	1	e_3^-	-1	e_3^+	-1	e_3^+	e_3^-
χ_{14}	9	$-9e_3^-$	$-9e_3^+$	$-e_3^-$	$-e_3^+$	1	0	0	$-e_3^-$	$-e_3^+$	1	e_3^+	-1	e_3^-	-1	e_3^-	e_3^+
χ_{15}	10	10	10	-2	-2	-2	1	1	0	0	0	0	0	0	0	0	0
χ_{16}	15	$-15e_3^+$	$-15e_3^-$	e_3^+	e_3^-	-1	0	0	e_3^+	e_3^-	-1	0	0	0	0	0	0
χ_{17}	15	$-15e_3^-$	$-15e_3^+$	e_3^-	e_3^+	-1	0	0	e_3^-	e_3^+	-1	0	0	0	0	0	0

TABLE 6. Character table for the centralizer of class 3A. Here, $\alpha = -e^{2\pi i \frac{2}{15}} - e^{2\pi i \frac{8}{15}}$ and $\beta = -e^{2\pi i \frac{1}{15}} - e^{2\pi i \frac{4}{15}}$.

Class 3B

$C_{M_{24}}(3B) \cong \mathbb{Z}_3 \times PSL_3(2)$ Order: $2^3 \cdot 3^2 \cdot 7 = 504$ Linear rep.

	504	504	504	9	9	9	21	21	21	21	21	21	24	24	24	12	12	12
	1A$_1$	3B$_1$	3B$_2$	3A$_1$	3B$_3$	3B$_4$	7A$_1$	21A$_1$	21A$_2$	7B$_1$	21B$_1$	21B$_2$	2B$_1$	6B$_1$	6B$_2$	4C$_1$	12B$_1$	12B$_2$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$
χ_3	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$
χ_4	3	3	3	0	0	0	$-e_7^+$	$-e_7^+$	$-e_7^+$	$-e_7^-$	$-e_7^-$	$-e_7^-$	-1	-1	-1	1	1	1
χ_5	3	3	3	0	0	0	$-e_7^-$	$-e_7^-$	$-e_7^-$	$-e_7^+$	$-e_7^+$	$-e_7^+$	-1	-1	-1	1	1	1
χ_6	3	$-3e_3^+$	$-3e_3^-$	0	0	0	$-e_7^+$	α	β	$-e_7^-$	$\bar{\beta}$	$\bar{\alpha}$	-1	e_3^+	e_3^-	1	$-e_3^+$	$-e_3^-$
χ_7	3	$-3e_3^+$	$-3e_3^-$	0	0	0	$-e_7^-$	$\bar{\beta}$	$\bar{\alpha}$	$-e_7^+$	α	β	-1	e_3^+	e_3^-	1	$-e_3^+$	$-e_3^-$
χ_8	3	$-3e_3^-$	$-3e_3^+$	0	0	0	$-e_7^+$	β	α	$-e_7^-$	$\bar{\alpha}$	$\bar{\beta}$	-1	e_3^-	e_3^+	1	$-e_3^-$	$-e_3^+$
χ_9	3	$-3e_3^-$	$-3e_3^+$	0	0	0	$-e_7^-$	$\bar{\alpha}$	$\bar{\beta}$	$-e_7^+$	β	α	-1	e_3^-	e_3^+	1	$-e_3^-$	$-e_3^+$
χ_{10}	6	6	6	0	0	0	-1	-1	-1	-1	-1	-1	2	2	2	0	0	0
χ_{11}	6	$-6e_3^+$	$-6e_3^-$	0	0	0	-1	e_3^+	e_3^-	-1	e_3^+	e_3^-	2	$-2e_3^+$	$-2e_3^-$	0	0	0
χ_{12}	6	$-6e_3^-$	$-6e_3^+$	0	0	0	-1	e_3^-	e_3^+	-1	e_3^-	e_3^+	2	$-2e_3^-$	$-2e_3^+$	0	0	0
χ_{13}	7	7	7	1	1	1	0	0	0	0	0	0	-1	-1	-1	-1	-1	-1
χ_{14}	7	$-7e_3^+$	$-7e_3^-$	1	$-e_3^+$	$-e_3^-$	0	0	0	0	0	0	-1	e_3^+	e_3^-	-1	e_3^+	e_3^-
χ_{15}	7	$-7e_3^-$	$-7e_3^+$	1	$-e_3^-$	$-e_3^+$	0	0	0	0	0	0	-1	e_3^-	e_3^+	-1	e_3^-	e_3^+
χ_{16}	8	8	8	-1	-1	-1	1	1	1	1	1	1	0	0	0	0	0	0
χ_{17}	8	$-8e_3^+$	$-8e_3^-$	-1	e_3^+	e_3^-	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	0	0	0	0	0	0
χ_{18}	8	$-8e_3^-$	$-8e_3^+$	-1	e_3^-	e_3^+	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	0	0	0	0	0	0

TABLE 7. Character table for the centralizer of class 3B. Here, $\alpha = e^{2\pi i \frac{2}{21}} + e^{2\pi i \frac{8}{21}} + e^{2\pi i \frac{11}{21}}$ and $\beta = e^{2\pi i \frac{1}{21}} + e^{2\pi i \frac{4}{21}} + e^{2\pi i \frac{16}{21}}$.

Class 4A

$C_{M_{24}}(4A)$ Order: $2^7 \cdot 3 = 384$

Genuinely projective rep. Central extension: $\mathbb{Z}_2.C_{M_{24}}(4A)$

	384	32	64	16	12	64	64	32	16	16	12	32	64	16	32	12	16	12	32	16	64	32	384	64	384	384
	1A$_1$	2A$_1$	2A$_2$	4B$_1$	3A$_1$	2B$_3$	2A$_3$	2B$_2$	4C$_1$	4A$_1$	6A$_1$	4A$_2$	4A$_3$	4B$_2$	2B$_1$	12A$_1$	4C$_2$	12A$_2$	4B$_3$	4A$_4$	4A$_5$	4B$_4$	4A$_6$	4A$_7$	4A$_8$	2A$_4$
χ_1	2	0	0	$\sqrt{2}$	-1	-2	0	0	0	$i\sqrt{2}$	1	0	0	$\sqrt{2}$	0	i	0	$-i$	0	$i\sqrt{2}$	$-2i$	0	$-2i$	0	$2i$	-2
χ_2	2	0	0	$-\sqrt{2}$	-1	-2	0	0	0	$-i\sqrt{2}$	1	0	0	$-\sqrt{2}$	0	i	0	$-i$	0	$-i\sqrt{2}$	$-2i$	0	$-2i$	0	$2i$	-2
χ_3	2	0	0	$\sqrt{2}$	-1	-2	0	0	0	$i\sqrt{2}$	1	0	0	$-\sqrt{2}$	0	$-i$	0	i	0	$-i\sqrt{2}$	$2i$	0	$2i$	0	$-2i$	-2
χ_4	2	0	0	$-\sqrt{2}$	-1	-2	0	0	0	$-i\sqrt{2}$	1	0	0	$\sqrt{2}$	0	$-i$	0	i	0	$i\sqrt{2}$	$2i$	0	$2i$	0	$-2i$	-2
χ_5	4	0	0	0	1	-4	0	0	0	0	-1	0	0	0	0	$-i$	0	i	0	0	$-4i$	0	$-4i$	0	$4i$	-4
χ_6	4	0	0	0	1	-4	0	0	0	0	-1	0	0	0	0	i	0	$-i$	0	0	$4i$	0	$4i$	0	$-4i$	-4
χ_7	4	0	0	0	1	0	0	0	-2	0	1	0	0	0	0	-1	$-2i$	-1	0	0	0	0	-4	0	-4	4
χ_8	4	0	0	0	1	0	0	0	-2	0	1	0	0	0	0	1	$2i$	1	0	0	0	0	4	0	4	4
χ_9	6	0	0	$-\sqrt{2}$	0	2	0	0	0	$i\sqrt{2}$	0	0	0	$\sqrt{2}$	0	0	0	0	0	$-i\sqrt{2}$	$2i$	0	$-6i$	0	$6i$	-6
χ_{10}	6	0	0	$\sqrt{2}$	0	2	0	0	0	$-i\sqrt{2}$	0	0	0	$-\sqrt{2}$	0	0	0	0	0	$i\sqrt{2}$	$2i$	0	$-6i$	0	$6i$	-6
χ_{11}	4	0	0	0	1	0	0	0	2	0	1	0	0	0	0	-1	$2i$	-1	0	0	0	0	-4	0	-4	4
χ_{12}	4	0	0	0	1	0	0	0	2	0	1	0	0	0	0	1	$-2i$	1	0	0	0	0	4	0	4	4
χ_{13}	6	0	0	$-\sqrt{2}$	0	2	0	0	0	$i\sqrt{2}$	0	0	0	$-\sqrt{2}$	0	0	0	0	0	$i\sqrt{2}$	$-2i$	0	$6i$	0	$-6i$	-6
χ_{14}	6	0	0	$\sqrt{2}$	0	2	0	0	0	$-i\sqrt{2}$	0	0	0	$\sqrt{2}$	0	0	0	0	0	$-i\sqrt{2}$	$-2i$	0	$6i$	0	$-6i$	-6
χ_{15}	8	0	0	0	-1	0	0	0	0	0	-1	0	0	0	0	-1	0	-1	0	0	0	0	8	0	8	8
χ_{16}	8	0	0	0	-1	0	0	0	0	0	-1	0	0	0	0	1	0	1	0	0	0	0	-8	0	-8	8

TABLE 8. Projective character table for the centralizer of class 4A.

Class 4B

$C_{M_{24}}(4B) \cong ((\mathbb{Z}_4 \times \mathbb{Z}_4) \rtimes \mathbb{Z}_4) \rtimes \mathbb{Z}_2$ Order: 128

Genuinely projective rep. Central extension: $\mathbb{Z}_2.C_{M_{24}}(4B)$

	128	16	128	128	64	64	128	16	16	16	16	32	16	16	32	32	32	16	16	16	16	64	16	64	64	64
	1A	8A$_1$	4B$_2$	4B$_6$	2B$_3$	4B$_1$	2A$_1$	2B$_1$	4B$_3$	2A$_3$	4B$_5$	2B$_2$	8A$_2$	8A$_3$	4A$_2$	4A$_1$	2A$_4$	4B$_7$	4B$_4$	4A$_3$	4A$_4$	2A$_2$	8A$_4$	4B$_9$	2A$_5$	4B$_8$
χ_1	2	-2	-2	-2	2	2	2	0	0	0	0	0	0	0	0	0	0	$\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$	$-\sqrt{2}$	0	0	0	0	0
χ_2	2	-2	-2	-2	2	2	2	0	0	0	0	0	0	0	0	0	0	$-\sqrt{2}$	$-\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	0	0	0	0	0
χ_3	2	0	-2	-2	-2	-2	2	0	0	0	0	0	0	0	0	0	0	$\sqrt{2}$	$-\sqrt{2}$	$-\sqrt{2}$	$\sqrt{2}$	0	2	0	0	0
χ_4	2	0	-2	-2	-2	-2	2	0	0	0	0	0	0	0	0	0	0	$-\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$	0	2	0	0	0
χ_5	2	0	-2	-2	-2	-2	2	0	0	0	0	0	0	0	0	0	0	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{2}$	0	-2	0	0	0
χ_6	2	0	-2	-2	-2	-2	2	0	0	0	0	0	0	0	0	0	0	$-\sqrt{2}$	$-\sqrt{2}$	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	0	0	0
χ_7	2	2	-2	-2	2	2	2	0	0	0	0	0	0	0	0	0	0	$\sqrt{2}$	$-\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$	0	0	0	0	0
χ_8	2	2	-2	-2	2	2	2	0	0	0	0	0	0	0	0	0	0	$-\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$	$\sqrt{2}$	0	0	0	0	0
χ_9	4	0	4	4	-4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
χ_{10}	4	0	4	4	4	-4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
χ_{11}	4	0	$4i$	$-4i$	0	0	-4	0	0	0	0	0	$2i$	2	0	0	0	0	0	0	0	0	0	0	0	0
χ_{12}	4	0	$4i$	$-4i$	0	0	-4	0	0	0	0	0	$-2i$	-2	0	0	0	0	0	0	0	0	0	0	0	0
χ_{13}	4	0	$-4i$	$4i$	0	0	-4	0	0	0	0	0	2	$2i$	0	0	0	0	0	0	0	0	0	0	0	0
χ_{14}	4	0	$-4i$	$4i$	0	0	-4	0	0	0	0	0	-2	$-2i$	0	0	0	0	0	0	0	0	0	0	0	0

TABLE 9. Projective character table for the centralizer of class 4B

Class 4C

$C_{M_{24}}(4C) \cong \mathbb{Z}_4 \times S_4$ Order: $2^5 \cdot 3 = 96$

Genuinely projective rep. Central extension: $\mathbb{Z}_2.C_{M_{24}}(4C) \cong \mathbb{Z}_4 \times \mathbb{Z}_2.S_4$

	96	16	12	16	32	96	16	12	16	32	96	16	12	16	32	96	16	12	16	32
	1A	$2A_1$	$3B_1$	$4A_1$	$2A_2$	$4C_1$	$4C_3$	$12B_1$	$4C_7$	$4C_4$	$4C_2$	$4C_5$	$12B_2$	$4C_8$	$4C_6$	$2B_1$	$2B_2$	$6B_1$	$4A_2$	$2B_3$
χ_1	2	0	-1	$\sqrt{2}$	0	2	0	-1	$i\sqrt{2}$	0	2	0	-1	$i\sqrt{2}$	0	2	0	-1	$\sqrt{2}$	0
χ_2	2	0	-1	$-\sqrt{2}$	0	2	0	-1	$-i\sqrt{2}$	0	2	0	-1	$-i\sqrt{2}$	0	2	0	-1	$-\sqrt{2}$	0
χ_3	2	0	-1	$\sqrt{2}$	0	-2	0	1	$-i\sqrt{2}$	0	-2	0	1	$-i\sqrt{2}$	0	2	0	-1	$\sqrt{2}$	0
χ_4	2	0	-1	$-\sqrt{2}$	0	-2	0	1	$i\sqrt{2}$	0	-2	0	1	$i\sqrt{2}$	0	2	0	-1	$-\sqrt{2}$	0
χ_5	2	0	-1	$\sqrt{2}$	0	$-2i$	0	i	$\sqrt{2}$	0	$2i$	0	$-i$	$-\sqrt{2}$	0	-2	0	1	$-\sqrt{2}$	0
χ_6	2	0	-1	$-\sqrt{2}$	0	$-2i$	0	i	$-\sqrt{2}$	0	$2i$	0	$-i$	$\sqrt{2}$	0	-2	0	1	$\sqrt{2}$	0
χ_7	2	0	-1	$\sqrt{2}$	0	$2i$	0	$-i$	$-\sqrt{2}$	0	$-2i$	0	i	$\sqrt{2}$	0	-2	0	1	$-\sqrt{2}$	0
χ_8	2	0	-1	$-\sqrt{2}$	0	$2i$	0	$-i$	$\sqrt{2}$	0	$-2i$	0	i	$-\sqrt{2}$	0	-2	0	1	$\sqrt{2}$	0
χ_9	4	0	1	0	0	4	0	1	0	0	4	0	1	0	0	4	0	1	0	0
χ_{10}	4	0	1	0	0	-4	0	-1	0	0	-4	0	-1	0	0	4	0	1	0	0
χ_{11}	4	0	1	0	0	$-4i$	0	$-i$	0	0	$4i$	0	i	0	0	-4	0	-1	0	0
χ_{12}	4	0	1	0	0	$4i$	0	i	0	0	$-4i$	0	$-i$	0	0	-4	0	-1	0	0

TABLE 10. Projective character table for the centralizer of class 4C.

Class 5A

$C_{M_{24}}(5A) \cong \mathbb{Z}_5 \times A_4$ Order: $2^2 \cdot 3 \cdot 5 = 60$ Linear rep.

	60	15	15	20	15	15	15	15	20	20	15	15	60	20	20	15	15	60	60	60
	1A	$3A_1$	$3A_2$	$10A_1$	$15A_1$	$15B_1$	$15A_2$	$15B_2$	$10A_2$	$10A_3$	$15B_3$	$15A_3$	$5A_1$	$2B_1$	$10A_4$	$15A_4$	$15B_4$	$5A_2$	$5A_3$	$5A_4$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	1	1	ε_5^4	ε_5^4	ε_5^4	ε_5^2	ε_5^2	ε_5^2	ε_5	ε_5	ε_5	ε_5^2	1	ε_5^3	ε_5^3	ε_5^3	ε_5^4	ε_5^3	ε_5
χ_3	1	1	1	ε_5^3	ε_5^3	ε_5^3	ε_5^4	ε_5^4	ε_5^4	ε_5^2	ε_5^2	ε_5^2	ε_5^4	1	ε_5	ε_5	ε_5	ε_5^3	ε_5	ε_5^2
χ_4	1	1	1	ε_5^2	ε_5^2	ε_5^2	ε_5	ε_5	ε_5	ε_5^3	ε_5^3	ε_5^3	ε_5	1	ε_5^4	ε_5^4	ε_5^4	ε_5^2	ε_5^4	ε_5^3
χ_5	1	1	1	ε_5	ε_5	ε_5	ε_5^3	ε_5^3	ε_5^3	ε_5^4	ε_5^4	ε_5^4	ε_5^3	1	ε_5^2	ε_5^2	ε_5^2	ε_5	ε_5^2	ε_5^4
χ_6	1	$-e_3^+$	$-e_3^-$	1	$-e_3^+$	$-e_3^-$	$-e_3^-$	$-e_3^+$	1	1	$-e_3^-$	$-e_3^+$	1	1	1	$-e_3^-$	$-e_3^+$	1	1	1
χ_7	1	$-e_3^-$	$-e_3^+$	1	$-e_3^-$	$-e_3^+$	$-e_3^+$	$-e_3^-$	1	1	$-e_3^+$	$-e_3^-$	1	1	1	$-e_3^+$	$-e_3^-$	1	1	1
χ_8	1	$-e_3^+$	$-e_3^-$	ε_5^4	ε_{15}^7	ε_{15}^2	ε_{15}^{11}	ε_{15}	ε_5^2	ε_5	ε_{15}^8	ε_{15}^{13}	ε_5^2	1	ε_5^3	ε_{15}^{14}	ε_{15}^4	ε_5^4	ε_5^3	ε_5
χ_9	1	$-e_3^+$	$-e_3^-$	ε_5^3	ε_{15}^4	ε_{15}^{14}	ε_{15}^2	ε_{15}^7	ε_5^4	ε_5^2	ε_{15}^{11}	ε_{15}	ε_5^4	1	ε_5	ε_{15}^8	ε_{15}^{13}	ε_5^3	ε_5	ε_5^2
χ_{10}	1	$-e_3^+$	$-e_3^-$	ε_5^2	ε_{15}	ε_{15}^{11}	ε_{15}^8	ε_{15}^{13}	ε_5	ε_5^3	ε_{15}^{14}	ε_{15}^4	ε_5	1	ε_5^4	ε_{15}^2	ε_{15}^7	ε_5^2	ε_5^4	ε_5^3
χ_{11}	1	$-e_3^+$	$-e_3^-$	ε_5	ε_{15}^{13}	ε_{15}^8	ε_{15}^{14}	ε_{15}^4	ε_5^3	ε_5^4	ε_{15}^2	ε_{15}^7	ε_5^3	1	ε_5^2	ε_{15}^{11}	ε_{15}	ε_5	ε_5^2	ε_5^4
χ_{12}	1	$-e_3^-$	$-e_3^+$	ε_5^4	ε_{15}^2	ε_{15}^7	ε_{15}	ε_{15}^{11}	ε_5^2	ε_5	ε_{15}^{13}	ε_{15}^8	ε_5^2	1	ε_5^3	ε_{15}^4	ε_{15}^{14}	ε_5^4	ε_5^3	ε_5
χ_{13}	1	$-e_3^-$	$-e_3^+$	ε_5^3	ε_{15}^{14}	ε_{15}^4	ε_{15}^7	ε_{15}^2	ε_5^4	ε_5^2	ε_{15}	ε_{15}^{11}	ε_5^4	1	ε_5	ε_{15}^{13}	ε_{15}^8	ε_5^3	ε_5	ε_5^2
χ_{14}	1	$-e_3^-$	$-e_3^+$	ε_5^2	ε_{15}^{11}	ε_{15}	ε_{15}^{13}	ε_{15}^8	ε_5	ε_5^3	ε_{15}^4	ε_{15}^{14}	ε_5	1	ε_5^4	ε_{15}^7	ε_{15}^2	ε_5^2	ε_5^4	ε_5^3
χ_{15}	1	$-e_3^-$	$-e_3^+$	ε_5	ε_{15}^8	ε_{15}^{13}	ε_{15}^4	ε_{15}^{14}	ε_5^3	ε_5^4	ε_{15}^7	ε_{15}^2	ε_5^3	1	ε_5^2	ε_{15}	ε_{15}^{11}	ε_5	ε_5^2	ε_5^4
χ_{16}	3	0	0	-1	0	0	0	0	-1	-1	0	0	3	-1	-1	0	0	3	3	3
χ_{17}	3	0	0	$-\varepsilon_5^2$	0	0	0	0	$-\varepsilon_5$	$-\varepsilon_5^3$	0	0	$3\varepsilon_5$	-1	$-\varepsilon_5^4$	0	0	$3\varepsilon_5^2$	$3\varepsilon_5^4$	$3\varepsilon_5^3$
χ_{18}	3	0	0	$-\varepsilon_5$	0	0	0	0	$-\varepsilon_5^3$	$-\varepsilon_5^4$	0	0	$3\varepsilon_5^3$	-1	$-\varepsilon_5^2$	0	0	$3\varepsilon_5$	$3\varepsilon_5^2$	$3\varepsilon_5^4$
χ_{19}	3	0	0	$-\varepsilon_5^4$	0	0	0	0	$-\varepsilon_5^2$	$-\varepsilon_5$	0	0	$3\varepsilon_5^2$	-1	$-\varepsilon_5^3$	0	0	$3\varepsilon_5^4$	$3\varepsilon_5^3$	$3\varepsilon_5$
χ_{20}	3	0	0	$-\varepsilon_5^3$	0	0	0	0	$-\varepsilon_5^4$	$-\varepsilon_5^2$	0	0	$3\varepsilon_5^4$	-1	$-\varepsilon_5$	0	0	$3\varepsilon_5^3$	$3\varepsilon_5$	$3\varepsilon_5^2$

TABLE 11. Character table for the centralizer of class 5A.

Classes 6A and 6B

$C_{M_{24}}(6A) \cong C_{M_{24}}(6B) \cong \mathbb{Z}_6.\mathbb{Z}_2^2 \cong \mathbb{Z}_3 \times D_8$ Order: 24 Linear rep.

	24	12	12	24	24	12	12	12	12	12	12	12	24	24	24
$C_{M_{24}}(6A)$	1A	2A$_1$	6A$_1$	3A$_1$	3A$_2$	6A$_2$	2A$_2$	4A$_1$	12A$_1$	6A$_3$	12A$_2$	6A$_4$	6A$_5$	2A$_3$	6A$_6$
$C_{M_{24}}(6B)$	1A	2B$_1$	6B$_2$	3B$_2$	3B$_1$	6B$_5$	2B$_2$	4C$_1$	12B$_1$	6B$_3$	12B$_2$	6B$_1$	6B$_4$	2B$_3$	6B$_6$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	−1	−1	1	1	−1	−1	1	1	−1	1	−1	1	1	1
χ_3	1	−1	−1	1	1	−1	1	−1	−1	1	−1	1	1	1	1
χ_4	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	1	1	1
χ_5	1	−1	e_3^-	$-e_3^-$	$-e_3^+$	e_3^+	−1	1	$-e_3^+$	e_3^+	$-e_3^-$	e_3^-	$-e_3^+$	1	$-e_3^-$
χ_6	1	−1	e_3^+	$-e_3^+$	$-e_3^-$	e_3^-	−1	1	$-e_3^-$	e_3^-	$-e_3^+$	e_3^+	$-e_3^-$	1	$-e_3^+$
χ_7	1	−1	e_3^-	$-e_3^-$	$-e_3^+$	e_3^+	1	−1	e_3^+	$-e_3^+$	e_3^-	$-e_3^-$	$-e_3^+$	1	$-e_3^-$
χ_8	1	−1	e_3^+	$-e_3^+$	$-e_3^-$	e_3^-	1	−1	e_3^-	$-e_3^-$	e_3^+	$-e_3^+$	$-e_3^-$	1	$-e_3^+$
χ_9	1	1	$-e_3^+$	$-e_3^+$	$-e_3^-$	$-e_3^-$	−1	−1	e_3^-	e_3^-	e_3^+	e_3^+	$-e_3^-$	1	$-e_3^+$
χ_{10}	1	1	$-e_3^-$	$-e_3^-$	$-e_3^+$	$-e_3^+$	−1	−1	e_3^+	e_3^+	e_3^-	e_3^-	$-e_3^+$	1	$-e_3^-$
χ_{11}	1	1	$-e_3^+$	$-e_3^+$	$-e_3^-$	$-e_3^-$	1	1	$-e_3^-$	$-e_3^-$	$-e_3^+$	$-e_3^+$	$-e_3^-$	1	$-e_3^+$
χ_{12}	1	1	$-e_3^-$	$-e_3^-$	$-e_3^+$	$-e_3^+$	1	1	$-e_3^+$	$-e_3^+$	$-e_3^-$	$-e_3^-$	$-e_3^+$	1	$-e_3^-$
χ_{13}	2	0	0	2	2	0	0	0	0	0	0	0	−2	−2	−2
χ_{14}	2	0	0	$-2e_3^+$	$-2e_3^-$	0	0	0	0	0	0	0	$2e_3^-$	−2	$2e_3^+$
χ_{15}	2	0	0	$-2e_3^-$	$-2e_3^+$	0	0	0	0	0	0	0	$2e_3^+$	−2	$2e_3^-$

Table 12. Character table for the centralizer of classes 6AB. NOTE: There is an accidental isomorphism $C_{M_{24}}(6A) \cong C_{M_{24}}(6B)$, therefore the character table is given only once. However, following our conventions, we give different names to the conjugacy classes depending on whether the group is $C_{M_{24}}(6A)$ or $C_{M_{24}}(6B)$.

Classes 7A and 7B

$C_{M_{24}}(7A) \cong C_{M_{24}}(7B) \cong S_3 \times \mathbb{Z}_7$ Order: $2 \cdot 3 \cdot 7 = 42$ Linear rep.

	42	14	14	21	42	14	21	14	21	21	14	14	42	42	42	42	14	21	21	42	21
$C_{M_{24}}$(7A)	1A	2A$_1$	14B$_1$	21B$_1$	7B$_1$	14B$_2$	21B$_2$	14A$_1$	21A$_1$	21A$_2$	14A$_2$	14A$_3$	7A$_1$	7A$_2$	7A$_3$	7B$_2$	14B$_3$	3B$_1$	21B$_3$	7B$_3$	21A$_3$
$C_{M_{24}}$(7B)	1A	2A$_1$	14A$_1$	21A$_1$	7A$_1$	14A$_2$	21A$_2$	14B$_1$	21B$_1$	21B$_2$	14B$_2$	14B$_3$	7B$_1$	7B$_2$	7B$_3$	7A$_2$	14A$_3$	3B$_1$	21A$_3$	7A$_3$	21B$_3$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	−1	−1	1	1	−1	1	−1	1	1	−1	−1	1	1	1	1	−1	1	1	1	1
χ_3	1	−1	$-\varepsilon_7$	ε_7	ε_7	$-\varepsilon_7^4$	ε_7^4	$-\varepsilon_7^3$	ε_7^3	ε_7^6	$-\varepsilon_7^6$	$-\varepsilon_7^5$	ε_7^5	ε_7^3	ε_7^6	ε_7^2	$-\varepsilon_7^2$	1	ε_7^2	ε_7^4	ε_7^5
χ_4	1	−1	$-\varepsilon_7^2$	ε_7^2	ε_7^2	$-\varepsilon_7$	ε_7	$-\varepsilon_7^6$	ε_7^6	ε_7^5	$-\varepsilon_7^5$	$-\varepsilon_7^3$	ε_7^3	ε_7^6	ε_7^5	ε_7^4	$-\varepsilon_7^4$	1	ε_7^4	ε_7	ε_7^3
χ_5	1	−1	$-\varepsilon_7^3$	ε_7^3	ε_7^3	$-\varepsilon_7^5$	ε_7^5	$-\varepsilon_7^2$	ε_7^2	ε_7^4	$-\varepsilon_7^4$	$-\varepsilon_7$	ε_7	ε_7^2	ε_7^4	ε_7^6	$-\varepsilon_7^6$	1	ε_7^6	ε_7^5	ε_7
χ_6	1	−1	$-\varepsilon_7^4$	ε_7^4	ε_7^4	$-\varepsilon_7^2$	ε_7^2	$-\varepsilon_7^5$	ε_7^5	ε_7^3	$-\varepsilon_7^3$	$-\varepsilon_7^6$	ε_7^6	ε_7^5	ε_7^3	ε_7	$-\varepsilon_7$	1	ε_7	ε_7^2	ε_7^6
χ_7	1	−1	$-\varepsilon_7^5$	ε_7^5	ε_7^5	$-\varepsilon_7^6$	ε_7^6	$-\varepsilon_7$	ε_7	ε_7^2	$-\varepsilon_7^2$	$-\varepsilon_7^4$	ε_7^4	ε_7	ε_7^2	ε_7^3	$-\varepsilon_7^3$	1	ε_7^3	ε_7^6	ε_7^4
χ_8	1	−1	$-\varepsilon_7^6$	ε_7^6	ε_7^6	$-\varepsilon_7^3$	ε_7^3	$-\varepsilon_7^4$	ε_7^4	ε_7	$-\varepsilon_7$	$-\varepsilon_7^2$	ε_7^2	ε_7^4	ε_7	ε_7^5	$-\varepsilon_7^5$	1	ε_7^5	ε_7^3	ε_7^2
χ_9	1	1	ε_7^6	ε_7^6	ε_7^6	ε_7^3	ε_7^3	ε_7^4	ε_7^4	ε_7	ε_7	ε_7^2	ε_7^2	ε_7^4	ε_7	ε_7^5	ε_7^5	1	ε_7^5	ε_7^3	ε_7^2
χ_{10}	1	1	ε_7^5	ε_7^5	ε_7^5	ε_7^6	ε_7^6	ε_7	ε_7	ε_7^2	ε_7^2	ε_7^4	ε_7^4	ε_7	ε_7^2	ε_7^3	ε_7^3	1	ε_7^3	ε_7^6	ε_7^4
χ_{11}	1	1	ε_7^4	ε_7^4	ε_7^4	ε_7^2	ε_7^2	ε_7^5	ε_7^5	ε_7^3	ε_7^3	ε_7^6	ε_7^6	ε_7^5	ε_7^3	ε_7	ε_7	1	ε_7	ε_7^2	ε_7^6
χ_{12}	1	1	ε_7^3	ε_7^3	ε_7^3	ε_7^5	ε_7^5	ε_7^2	ε_7^2	ε_7^4	ε_7^4	ε_7	ε_7	ε_7^2	ε_7^4	ε_7^6	ε_7^6	1	ε_7^6	ε_7^5	ε_7
χ_{13}	1	1	ε_7^2	ε_7^2	ε_7^2	ε_7	ε_7	ε_7^6	ε_7^6	ε_7^5	ε_7^5	ε_7^3	ε_7^3	ε_7^6	ε_7^5	ε_7^4	ε_7^4	1	ε_7^4	ε_7	ε_7^3
χ_{14}	1	1	ε_7	ε_7	ε_7	ε_7^4	ε_7^4	ε_7^3	ε_7^3	ε_7^6	ε_7^6	ε_7^5	ε_7^5	ε_7^3	ε_7^6	ε_7^2	ε_7^2	1	ε_7^2	ε_7^4	ε_7^5
χ_{15}	2	0	0	−1	2	0	−1	0	−1	−1	0	0	2	2	2	2	0	−1	−1	2	−1
χ_{16}	2	0	0	$-\varepsilon_7^3$	$2\varepsilon_7^3$	0	$-\varepsilon_7^5$	0	$-\varepsilon_7^2$	$-\varepsilon_7^4$	0	0	$2\varepsilon_7$	$2\varepsilon_7^2$	$2\varepsilon_7^4$	$2\varepsilon_7^6$	0	−1	$-\varepsilon_7^6$	$2\varepsilon_7^5$	$-\varepsilon_7$
χ_{17}	2	0	0	$-\varepsilon_7^2$	$2\varepsilon_7^2$	0	$-\varepsilon_7$	0	$-\varepsilon_7^6$	$-\varepsilon_7^5$	0	0	$2\varepsilon_7^3$	$2\varepsilon_7^6$	$2\varepsilon_7^5$	$2\varepsilon_7^4$	0	−1	$-\varepsilon_7^4$	$2\varepsilon_7$	$-\varepsilon_7^3$
χ_{18}	2	0	0	$-\varepsilon_7$	$2\varepsilon_7$	0	$-\varepsilon_7^4$	0	$-\varepsilon_7^3$	$-\varepsilon_7^6$	0	0	$2\varepsilon_7^5$	$2\varepsilon_7^3$	$2\varepsilon_7^6$	$2\varepsilon_7^2$	0	−1	$-\varepsilon_7^2$	$2\varepsilon_7^4$	$-\varepsilon_7^5$
χ_{19}	2	0	0	$-\varepsilon_7^6$	$2\varepsilon_7^6$	0	$-\varepsilon_7^3$	0	$-\varepsilon_7^4$	$-\varepsilon_7$	0	0	$2\varepsilon_7^2$	$2\varepsilon_7^4$	$2\varepsilon_7$	$2\varepsilon_7^5$	0	−1	$-\varepsilon_7^5$	$2\varepsilon_7^3$	$-\varepsilon_7^2$
χ_{20}	2	0	0	$-\varepsilon_7^5$	$2\varepsilon_7^5$	0	$-\varepsilon_7^6$	0	$-\varepsilon_7$	$-\varepsilon_7^2$	0	0	$2\varepsilon_7^4$	$2\varepsilon_7$	$2\varepsilon_7^2$	$2\varepsilon_7^3$	0	−1	$-\varepsilon_7^3$	$2\varepsilon_7^6$	$-\varepsilon_7^4$
χ_{21}	2	0	0	$-\varepsilon_7^4$	$2\varepsilon_7^4$	0	$-\varepsilon_7^2$	0	$-\varepsilon_7^5$	$-\varepsilon_7^3$	0	0	$2\varepsilon_7^6$	$2\varepsilon_7^5$	$2\varepsilon_7^3$	$2\varepsilon_7$	0	−1	$-\varepsilon_7$	$2\varepsilon_7^2$	$-\varepsilon_7^6$

TABLE 13. Character table for the centralizer of classes 7AB.

Class 8A

$C_{M_{24}}(8A) \cong \mathbb{Z}_2 \times \mathbb{Z}_8$ Order: $2^4 = 16$ Linear rep.

	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	1A	8A$_1$	8A$_2$	8A$_3$	4B$_1$	8A$_4$	4B$_2$	8A$_5$	8A$_6$	4B$_3$	8A$_7$	4B$_4$	8A$_8$	2A$_1$	2B$_1$	2B$_2$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	-1	-1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	1	1
χ_3	1	-1	1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1
χ_4	1	1	-1	1	-1	1	1	1	-1	-1	-1	1	-1	1	-1	-1
χ_5	1	$-i$	$-i$	i	-1	i	-1	$-i$	i	-1	i	-1	$-i$	1	1	1
χ_6	1	i	i	$-i$	-1	$-i$	-1	i	$-i$	-1	$-i$	-1	i	1	1	1
χ_7	1	$-i$	i	i	1	i	-1	$-i$	$-i$	1	$-i$	-1	i	1	-1	-1
χ_8	1	i	$-i$	$-i$	1	$-i$	-1	i	i	1	i	-1	$-i$	1	-1	-1
χ_9	1	$-\varepsilon_8$	$-\varepsilon_8$	ε_8^7	$-i$	$-\varepsilon_8^7$	i	ε_8	$-\varepsilon_8^7$	i	ε_8^7	$-i$	ε_8	-1	-1	1
χ_{10}	1	ε_8^7	ε_8^7	$-\varepsilon_8$	i	ε_8	$-i$	$-\varepsilon_8^7$	ε_8	$-i$	$-\varepsilon_8$	i	$-\varepsilon_8^7$	-1	-1	1
χ_{11}	1	$-\varepsilon_8^7$	$-\varepsilon_8^7$	ε_8	i	$-\varepsilon_8$	$-i$	ε_8^7	$-\varepsilon_8$	$-i$	ε_8	i	ε_8^7	-1	-1	1
χ_{12}	1	ε_8	ε_8	$-\varepsilon_8^7$	$-i$	ε_8^7	i	$-\varepsilon_8$	ε_8^7	i	$-\varepsilon_8^7$	$-i$	$-\varepsilon_8$	-1	-1	1
χ_{13}	1	$-\varepsilon_8$	ε_8	ε_8^7	i	$-\varepsilon_8^7$	i	ε_8	ε_8^7	$-i$	$-\varepsilon_8^7$	$-i$	$-\varepsilon_8$	-1	1	-1
χ_{14}	1	ε_8^7	$-\varepsilon_8^7$	$-\varepsilon_8$	$-i$	ε_8	$-i$	$-\varepsilon_8^7$	$-\varepsilon_8$	i	ε_8	i	ε_8^7	-1	1	-1
χ_{15}	1	$-\varepsilon_8^7$	ε_8^7	ε_8	$-i$	$-\varepsilon_8$	$-i$	ε_8^7	ε_8	i	$-\varepsilon_8$	i	$-\varepsilon_8^7$	-1	1	-1
χ_{16}	1	ε_8	$-\varepsilon_8$	$-\varepsilon_8^7$	i	ε_8^7	i	$-\varepsilon_8$	$-\varepsilon_8^7$	$-i$	ε_8^7	$-i$	ε_8	-1	1	-1

Table 14. Character table for the centralizer of class 8A.

Class 10A

$C_{M_{24}}(10A) \cong \mathbb{Z}_2 \times \mathbb{Z}_{10}$ Order: 20 Linear rep.

	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
	1A	$10A_1$	$10A_2$	$10A_3$	$5A_1$	$10A_4$	$2B_1$	$2B_2$	$10A_5$	$10A_6$	$10A_7$	$10A_8$	$5A_2$	$5A_3$	$10A_9$	$10A_{10}$	$2B_3$	$10A_{11}$	$5A_4$	$10A_{12}$
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_2	1	-1	-1	-1	1	1	1	-1	-1	1	-1	1	1	1	-1	-1	-1	-1	1	1
χ_3	1	1	1	1	1	-1	-1	-1	1	-1	-1	-1	1	1	-1	-1	1	-1	1	-1
χ_4	1	-1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	1	1	-1	1	1	-1
χ_5	1	ε_5^3	ε_5^4	ε_5^2	ε_5^4	ε_5^2	1	1	ε_5	ε_5^4	ε_5^2	ε_5^3	ε_5^3	ε_5	ε_5	ε_5^3	1	ε_5^4	ε_5^2	ε_5
χ_6	1	$-\varepsilon_5^3$	$-\varepsilon_5^4$	$-\varepsilon_5^2$	ε_5^4	ε_5^2	1	-1	$-\varepsilon_5$	ε_5^4	$-\varepsilon_5^2$	ε_5^3	ε_5^3	ε_5	$-\varepsilon_5$	$-\varepsilon_5^3$	-1	$-\varepsilon_5^4$	ε_5^2	ε_5
χ_7	1	ε_5^3	ε_5^4	ε_5^2	ε_5^4	$-\varepsilon_5^2$	-1	-1	ε_5	$-\varepsilon_5^4$	$-\varepsilon_5^2$	$-\varepsilon_5^3$	ε_5^3	ε_5	$-\varepsilon_5$	$-\varepsilon_5^3$	1	$-\varepsilon_5^4$	ε_5^2	$-\varepsilon_5$
χ_8	1	$-\varepsilon_5^3$	$-\varepsilon_5^4$	$-\varepsilon_5^2$	ε_5^4	$-\varepsilon_5^2$	-1	1	$-\varepsilon_5$	$-\varepsilon_5^4$	ε_5^2	$-\varepsilon_5^3$	ε_5^3	ε_5	ε_5	ε_5^3	-1	ε_5^4	ε_5^2	$-\varepsilon_5$
χ_9	1	ε_5	ε_5^3	ε_5^4	ε_5^3	ε_5^4	1	1	ε_5^2	ε_5^3	ε_5^4	ε_5	ε_5	ε_5^2	ε_5^2	ε_5	1	ε_5^3	ε_5^4	ε_5^2
χ_{10}	1	$-\varepsilon_5$	$-\varepsilon_5^3$	$-\varepsilon_5^4$	ε_5^3	ε_5^4	1	-1	$-\varepsilon_5^2$	ε_5^3	$-\varepsilon_5^4$	ε_5	ε_5	ε_5^2	$-\varepsilon_5^2$	$-\varepsilon_5$	-1	$-\varepsilon_5^3$	ε_5^4	ε_5^2
χ_{11}	1	ε_5	ε_5^3	ε_5^4	ε_5^3	$-\varepsilon_5^4$	-1	-1	ε_5^2	$-\varepsilon_5^3$	$-\varepsilon_5^4$	$-\varepsilon_5$	ε_5	ε_5^2	$-\varepsilon_5^2$	$-\varepsilon_5$	1	$-\varepsilon_5^3$	ε_5^4	$-\varepsilon_5^2$
χ_{12}	1	$-\varepsilon_5$	$-\varepsilon_5^3$	$-\varepsilon_5^4$	ε_5^3	$-\varepsilon_5^4$	-1	1	$-\varepsilon_5^2$	$-\varepsilon_5^3$	ε_5^4	$-\varepsilon_5$	ε_5	ε_5^2	ε_5^2	ε_5	-1	ε_5^3	ε_5^4	$-\varepsilon_5^2$
χ_{13}	1	ε_5^4	ε_5^2	ε_5	ε_5^2	ε_5	1	1	ε_5^3	ε_5^2	ε_5	ε_5^4	ε_5^4	ε_5^3	ε_5^3	ε_5^4	1	ε_5^2	ε_5	ε_5^3
χ_{14}	1	$-\varepsilon_5^4$	$-\varepsilon_5^2$	$-\varepsilon_5$	ε_5^2	ε_5	1	-1	$-\varepsilon_5^3$	ε_5^2	$-\varepsilon_5$	ε_5^4	ε_5^4	ε_5^3	$-\varepsilon_5^3$	$-\varepsilon_5^4$	-1	$-\varepsilon_5^2$	ε_5	ε_5^3
χ_{15}	1	ε_5^4	ε_5^2	ε_5	ε_5^2	$-\varepsilon_5$	-1	-1	ε_5^3	$-\varepsilon_5^2$	$-\varepsilon_5$	$-\varepsilon_5^4$	ε_5^4	ε_5^3	$-\varepsilon_5^3$	$-\varepsilon_5^4$	1	$-\varepsilon_5^2$	ε_5	$-\varepsilon_5^3$
χ_{16}	1	$-\varepsilon_5^4$	$-\varepsilon_5^2$	$-\varepsilon_5$	ε_5^2	$-\varepsilon_5$	-1	1	$-\varepsilon_5^3$	$-\varepsilon_5^2$	ε_5	$-\varepsilon_5^4$	ε_5^4	ε_5^3	ε_5^3	ε_5^4	-1	ε_5^2	ε_5	$-\varepsilon_5^3$
χ_{17}	1	ε_5^2	ε_5	ε_5^3	ε_5	ε_5^3	1	1	ε_5^4	ε_5	ε_5^3	ε_5^2	ε_5^2	ε_5^4	ε_5^4	ε_5^2	1	ε_5	ε_5^3	ε_5^4
χ_{18}	1	$-\varepsilon_5^2$	$-\varepsilon_5$	$-\varepsilon_5^3$	ε_5	ε_5^3	1	-1	$-\varepsilon_5^4$	ε_5	$-\varepsilon_5^3$	ε_5^2	ε_5^2	ε_5^4	$-\varepsilon_5^4$	$-\varepsilon_5^2$	-1	$-\varepsilon_5$	ε_5^3	ε_5^4
χ_{19}	1	ε_5^2	ε_5	ε_5^3	ε_5	$-\varepsilon_5^3$	-1	-1	ε_5^4	$-\varepsilon_5$	$-\varepsilon_5^3$	$-\varepsilon_5^2$	ε_5^2	ε_5^4	$-\varepsilon_5^4$	$-\varepsilon_5^2$	1	$-\varepsilon_5$	ε_5^3	$-\varepsilon_5^4$
χ_{20}	1	$-\varepsilon_5^2$	$-\varepsilon_5$	$-\varepsilon_5^3$	ε_5	$-\varepsilon_5^3$	-1	1	$-\varepsilon_5^4$	$-\varepsilon_5$	ε_5^3	$-\varepsilon_5^2$	ε_5^2	ε_5^4	ε_5^4	ε_5^2	-1	ε_5	ε_5^3	$-\varepsilon_5^4$

TABLE 15. Character table for the centralizer of class 10A.

Appendix E. Decompositions of twisted sectors

In this section we give the decomposition of the first 20 representation spaces $\mathcal{H}_g^n$ for all non-isomorphic twisted sectors. In each table, the integer in the i'th row and the r'th column is the multiplicity $h_{i,r}$ of the irreducible representation χ_i (numbered as in the corresponding character table) in the representation $\rho_{g,r}$ of eq. (6). When the centraliser $C_{M_{24}}(g)$ is a cyclic group $\mathbb{Z}_N$, $N = o(g)$ (classes 11A, 12A, 12B, 14AB, 15AB, 21AB, 23AB), the multiplicity $h_{i,r}$ of the (projective) irreducible characters χ_i (see eq. (180)) is only non-zero if $Nr + 1/\ell(g) \equiv i \mod N$, where $\ell(g)$ is the length of the shortest cycle of g as a permutation of 24 objects; then, only these values of $h_{i,r}$ are given.

i \ $2r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	0	0	0	0	0	0	0	4	0	2	0	14	0	18	0	52	0	96	0	204
2	0	0	0	0	0	0	0	0	0	4	0	6	0	14	0	34	0	72	0	132	0
3,4	0	1	0	0	0	0	0	1	0	2	0	6	0	18	0	32	0	72	0	137	0
5,6	0	0	1	0	0	0	2	0	4	0	16	0	28	0	72	0	141	0	306	0	570
7,8	0	0	0	0	0	2	0	2	0	10	0	18	0	50	0	96	0	216	0	408	0
9	0	0	0	0	2	0	2	0	12	0	22	0	68	0	130	0	304	0	578	0	1184
10	0	0	0	0	0	0	6	0	10	0	30	0	72	0	166	0	332	0	702	0	1346
11	0	0	0	2	0	2	0	6	0	20	0	48	0	106	0	240	0	486	0	972	0
12,13	0	0	0	0	0	2	0	8	0	18	0	48	0	108	0	236	0	486	0	978	0
14	0	0	0	0	2	0	4	0	14	0	36	0	80	0	184	0	392	0	788	0	1550

Table 16. Multiplicities of irreducible representations for the 2A-twisted sector.

i \ $4r$	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39
1	0	0	0	0	0	0	2	0	2	0	8	0	16	0	42	0	76	0	172	0
2,3	1	0	0	0	0	0	1	0	4	0	6	0	21	0	36	0	84	0	160	0
4	0	0	0	0	0	0	2	0	2	0	8	0	16	0	42	0	76	0	172	0
5,8	0	0	1	0	0	0	6	0	10	0	32	0	68	0	162	0	318	0	670	0
6,7	0	0	0	0	2	0	4	0	14	0	26	0	76	0	150	0	334	0	648	0
9,10	0	1	0	0	0	2	0	8	0	18	0	47	0	108	0	224	0	466	0	924
11,14	0	0	0	0	2	0	4	0	16	0	38	0	90	0	192	0	414	0	814	0
12,13	0	0	1	0	2	0	6	0	14	0	40	0	86	0	198	0	406	0	826	0
15,16	0	0	0	0	2	0	6	0	18	0	46	0	106	0	234	0	490	0	986	0
17	0	0	0	2	0	4	0	14	0	40	0	94	0	210	0	456	0	932	0	1836
18	0	0	0	2	0	8	0	22	0	58	0	140	0	318	0	682	0	1396	0	2760

Table 17. Multiplicities of irreducible representations for the 2B-twisted sector.

$3r$ i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0	0	0	0	0	0	0	0	2	0	0	2	0	0	10	0	0	16	0	0
2,3	0	1	0	0	0	0	0	2	0	0	6	0	0	20	0	0	41	0	0	104
4,5	0	0	0	0	0	0	1	0	0	6	0	0	10	0	0	34	0	0	74	0
6,7	0	0	0	0	0	0	0	0	6	0	0	14	0	0	42	0	0	92	0	0
8	0	0	0	0	2	0	0	2	0	0	16	0	0	30	0	0	94	0	0	194
9	1	0	0	0	0	0	4	0	0	6	0	0	28	0	0	60	0	0	162	0
10,11	0	0	0	0	0	2	0	0	8	0	0	24	0	0	64	0	0	154	0	0
12	0	0	2	0	0	2	0	0	10	0	0	24	0	0	74	0	0	172	0	0
13	0	0	0	0	2	0	0	6	0	0	20	0	0	50	0	0	134	0	0	298
14	0	0	0	0	0	0	4	0	0	12	0	0	40	0	0	94	0	0	236	0
15	0	0	0	0	0	4	0	0	8	0	0	32	0	0	76	0	0	200	0	0
16	0	0	0	0	2	0	0	10	0	0	30	0	0	88	0	0	216	0	0	510
17	0	0	0	2	0	0	6	0	0	22	0	0	62	0	0	164	0	0	382	0

TABLE 18. Multiplicities of irreducible representations for the 3A-twisted sector.

$9r$ i	2	5	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62
1	0	0	0	0	0	2	0	0	2	0	0	10	0	0	18	0	0	54	0	0	108
2	0	0	0	0	0	0	0	2	0	0	4	0	0	18	0	0	36	0	0	90	0
3	0	0	0	0	0	0	0	0	0	6	0	0	10	0	0	30	0	0	62	0	0
4,5	0	0	0	0	0	2	0	0	10	0	0	22	0	0	65	0	0	146	0	0	344
6,7	0	0	0	0	2	0	0	5	0	0	18	0	0	44	0	0	118	0	0	254	0
8,9	1	0	0	0	0	0	4	0	0	10	0	0	36	0	0	82	0	0	202	0	0
10	0	0	0	0	0	6	0	0	14	0	0	50	0	0	122	0	0	308	0	0	666
11	0	2	0	0	2	0	0	12	0	0	32	0	0	96	0	0	220	0	0	526	0
12	0	0	0	2	0	0	6	0	0	26	0	0	64	0	0	174	0	0	390	0	0
13	0	0	2	0	0	4	0	0	20	0	0	54	0	0	150	0	0	346	0	0	794
14	0	0	0	0	4	0	0	12	0	0	42	0	0	104	0	0	268	0	0	600	0
15	0	0	0	2	0	0	10	0	0	26	0	0	80	0	0	194	0	0	468	0	0
16	0	0	2	0	0	6	0	0	22	0	0	64	0	0	168	0	0	400	0	0	902
17	0	0	0	0	4	0	0	14	0	0	46	0	0	122	0	0	304	0	0	690	0
18	0	0	0	2	0	0	10	0	0	32	0	0	90	0	0	224	0	0	530	0	0

TABLE 19. Multiplicities of irreducible representations for the 3B-twisted sector.

i \ $8r$	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39
1,2	1	0	0	0	0	0	0	0	4	0	0	0	10	0	0	0	26	0	0	0
3,4	0	0	0	0	0	0	2	0	0	0	4	0	0	0	16	0	0	0	44	0
5	0	0	0	0	2	0	0	0	6	0	0	0	18	0	0	0	54	0	0	0
6	0	0	0	0	0	0	2	0	0	0	12	0	0	0	32	0	0	0	84	0
7,11	0	1	0	0	0	2	0	0	0	7	0	0	0	26	0	0	0	68	0	0
8,12	0	0	0	0	0	0	0	4	0	0	0	14	0	0	0	40	0	0	0	110
9,10	0	0	0	0	2	0	0	0	8	0	0	0	28	0	0	0	80	0	0	0
13	0	0	2	0	0	0	4	0	0	0	16	0	0	0	48	0	0	0	130	0
14	0	0	0	0	0	0	4	0	0	0	16	0	0	0	48	0	0	0	128	0
15	0	0	0	2	0	0	0	8	0	0	0	28	0	0	0	84	0	0	0	216
16	0	0	0	0	0	4	0	0	0	16	0	0	0	48	0	0	0	136	0	0

Table 20. Multiplicities of irreducible representations for the 4A-twisted sector.

i \ $4r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1,6	0	1	0	0	0	3	0	0	0	10	0	0	0	33	0	0	0	90	0	0
2,3,4,5,7,8	0	0	0	0	0	2	0	0	0	10	0	0	0	32	0	0	0	90	0	0
9,10	0	0	0	2	0	0	0	10	0	0	0	36	0	0	0	110	0	0	0	290
11	1	0	0	0	2	0	0	0	16	0	0	0	46	0	0	0	146	0	0	0
12	0	0	0	0	4	0	0	0	12	0	0	0	52	0	0	0	136	0	0	0
13	0	0	0	0	0	0	8	0	0	0	24	0	0	0	88	0	0	0	224	0
14	0	0	2	0	0	0	6	0	0	0	30	0	0	0	80	0	0	0	236	0

Table 21. Multiplicities of irreducible representations for the 4B-twisted sector.

i \ $16r$	3	7	11	15	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79
1,2	0	0	0	2	0	0	0	8	0	0	0	30	0	0	0	90	0	0	0	228
3,4	0	0	0	0	0	4	0	0	0	18	0	0	0	52	0	0	0	144	0	0
5,6	0	0	1	0	0	0	7	0	0	0	22	0	0	0	68	0	0	0	184	0
7,8	1	0	0	0	2	0	0	0	12	0	0	0	41	0	0	0	112	0	0	0
9	0	0	0	4	0	0	0	18	0	0	0	60	0	0	0	176	0	0	0	460
10	0	2	0	0	0	8	0	0	0	32	0	0	0	106	0	0	0	288	0	0
11	0	0	2	0	0	0	12	0	0	0	46	0	0	0	136	0	0	0	364	0
12	0	0	0	0	6	0	0	0	24	0	0	0	80	0	0	0	228	0	0	0

Table 22. Multiplicities of irreducible representations for the 4C-twisted sector.

i \ $5r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	20	0	0	0	0	60	0
2	0	0	0	0	0	0	0	2	0	0	0	0	14	0	0	0	0	42	0	0	0
3	1	0	0	0	0	2	0	0	0	0	6	0	0	0	0	28	0	0	0	0	74
4	0	0	0	2	0	0	0	0	4	0	0	0	0	16	0	0	0	0	50	0	0
5	0	0	0	0	0	0	2	0	0	0	0	10	0	0	0	0	30	0	0	0	0
6,7	0	0	0	0	1	0	0	0	0	6	0	0	0	0	20	0	0	0	0	63	0
8,12	0	0	0	0	0	0	0	4	0	0	0	0	12	0	0	0	0	42	0	0	0
9,13	0	0	0	0	0	2	0	0	0	0	7	0	0	0	0	26	0	0	0	0	74
10,14	0	0	0	0	0	0	0	0	4	0	0	0	0	18	0	0	0	0	48	0	0
11,15	0	1	0	0	0	0	2	0	0	0	0	10	0	0	0	0	32	0	0	0	0
16	0	0	0	0	4	0	0	0	0	16	0	0	0	0	64	0	0	0	0	180	0
17	0	0	0	2	0	0	0	0	14	0	0	0	0	46	0	0	0	0	154	0	0
18	0	0	0	0	0	0	8	0	0	0	0	28	0	0	0	0	100	0	0	0	0
19	0	0	2	0	0	0	0	8	0	0	0	0	40	0	0	0	0	118	0	0	0
20	1	0	0	0	0	4	0	0	0	0	24	0	0	0	0	76	0	0	0	0	230

TABLE 23. Multiplicities of irreducible representations for the 5A-twisted sector.

i \ $6r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1,2	0	0	0	0	0	4	0	0	0	0	0	14	0	0	0	0	0	56	0	0
3,4	0	0	0	0	0	2	0	0	0	0	0	16	0	0	0	0	0	52	0	0
5,12	0	1	0	0	0	0	0	5	0	0	0	0	0	25	0	0	0	0	0	76
6,11	0	0	0	1	0	0	0	0	0	10	0	0	0	0	0	34	0	0	0	0
7,10	0	0	0	0	0	0	0	6	0	0	0	0	0	22	0	0	0	0	0	80
8,9	0	0	0	2	0	0	0	0	0	8	0	0	0	0	0	38	0	0	0	0
13	0	0	2	0	0	0	0	0	14	0	0	0	0	0	58	0	0	0	0	0
14	1	0	0	0	0	0	8	0	0	0	0	0	38	0	0	0	0	0	130	0
15	0	0	0	0	4	0	0	0	0	0	24	0	0	0	0	0	88	0	0	0

TABLE 24. Multiplicities of irreducible representations for the 6A-twisted sector.

i \ $36r$	5	11	17	23	29	35	41	47	53	59	65	71	77	83	89	95	101	107	113	119	125	131	137	143
1,2	0	0	0	0	2	0	0	0	0	0	16	0	0	0	0	0	50	0	0	0	0	0	158	0
3,4	0	0	0	0	3	0	0	0	0	0	13	0	0	0	0	0	54	0	0	0	0	0	152	0
5,12	0	0	0	0	0	0	6	0	0	0	0	0	22	0	0	0	0	0	78	0	0	0	0	0
6,11	0	0	2	0	0	0	0	0	8	0	0	0	0	0	36	0	0	0	0	0	106	0	0	0
7,10	1	0	0	0	0	0	4	0	0	0	0	0	24	0	0	0	0	0	74	0	0	0	0	0
8,9	0	0	1	0	0	0	0	0	10	0	0	0	0	0	33	0	0	0	0	0	112	0	0	0
13	0	2	0	0	0	0	0	14	0	0	0	0	0	56	0	0	0	0	0	182	0	0	0	0
14	0	0	0	0	0	8	0	0	0	0	0	36	0	0	0	0	0	126	0	0	0	0	0	366
15	0	0	0	4	0	0	0	0	0	22	0	0	0	0	0	86	0	0	0	0	0	260	0	0

Table 25. Multiplicities of irreducible representations for the 6B-twisted sector.

i \ $7r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0	0	0	0	0	0	2	0	0	0	0	0	0	8	0	0	0	0	0	0
2	0	0	0	0	0	0	2	0	0	0	0	0	0	10	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	15	0	0	0
4	0	0	0	0	0	2	0	0	0	0	0	0	6	0	0	0	0	0	0	28
5	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	14	0	0	0	0
6	0	0	0	0	1	0	0	0	0	0	0	6	0	0	0	0	0	0	22	0
7	0	0	0	0	0	0	0	3	0	0	0	0	0	0	10	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	20	0	0
9	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	16	0	0
10	1	0	0	0	0	0	0	2	0	0	0	0	0	0	12	0	0	0	0	0
11	0	0	0	0	2	0	0	0	0	0	0	5	0	0	0	0	0	0	24	0
12	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	12	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	26
14	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	18	0	0	0
15	0	0	0	0	0	0	4	0	0	0	0	0	0	16	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	26	0	0	0	0
17	0	0	0	0	0	2	0	0	0	0	0	0	16	0	0	0	0	0	0	52
18	0	0	1	0	0	0	0	0	0	8	0	0	0	0	0	0	32	0	0	0
19	0	0	0	2	0	0	0	0	0	0	10	0	0	0	0	0	0	36	0	0
20	1	0	0	0	0	0	0	4	0	0	0	0	0	0	22	0	0	0	0	0
21	0	0	0	0	2	0	0	0	0	0	0	11	0	0	0	0	0	0	46	0

Table 26. Multiplicities of irreducible representations for the 7AB-twisted sectors.

i / $8r$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1,4	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	24	0	0	0	0
2,3	0	0	0	2	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	46
5	0	0	0	0	0	2	0	0	0	0	0	0	0	16	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	34	0	0
7	0	0	0	0	0	4	0	0	0	0	0	0	0	16	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	32	0	0
9,13	1	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	28	0	0	0
10,14	0	0	1	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	40	0
11,15	0	0	0	0	0	0	4	0	0	0	0	0	0	0	20	0	0	0	0	0
12,16	0	0	0	0	2	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0

Table 27. Multiplicities of irreducible representations for the 8A-twisted sector.

i / $20r$	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39	41
1,2	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	22	0
3,4	0	0	0	2	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
5,6	0	0	0	0	2	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0
7,8	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	24
9,10	1	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0
11,12	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	14	0	0	0	0
13,14	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	16	0	0	0
15,16	0	1	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0
17,18	0	0	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0
19,20	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	18	0	0

Table 28. Multiplicities of irreducible representations for the 10A-twisted sector.

$11r \equiv i-1$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$h_{i,r}$	2	0	2	2	2	2	4	4	6	6	8	8	12	10	14	16	20	20	26	26

Table 29. Multiplicities of irreducible representations for the 11A-twisted sector.

$24r \equiv 2i-1$	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39	41
$h_{i,r}$	1	2	2	2	2	4	4	4	6	6	8	8	10	12	12	16	18	20	22	24

Table 30. Multiplicities of irreducible representations for the 12A-twisted sector.

$144r \equiv 12i-1$	23	35	47	59	71	83	95	107	119	131	143	155	167	179	191	203	215	227	239	251
$h_{i,r}$	2	2	2	2	2	4	4	6	6	6	8	10	10	12	14	16	18	20	24	26

Table 31. Multiplicities of irreducible representations for the 12B-twisted sector.

$14r \equiv i-1$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$h_{i,r}$	1	0	1	2	1	2	2	2	3	4	4	4	6	6	7	8	8	10	12	12

Table 32. Multiplicities of irreducible representations for the 14AB-twisted sectors.

$15r \equiv i-1$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$h_{i,r}$	1	0	1	1	2	1	2	2	3	2	4	4	5	4	6	6	8	7	10	10

Table 33. Multiplicities of irreducible representations for the 15AB-twisted sectors.

$63r \equiv 3i-1$	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62	65
$h_{i,r}$	7	7	13	0	14	7	14	7	14	26	14	14	21	14	28	21	58	28	588	28

Table 34. Multiplicities of irreducible representations for the 21AB-twisted sectors.

$23r \equiv i-1$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$h_{i,r}$	1	0	0	1	1	0	1	0	2	1	2	2	2	1	2	2	2	2	3	2

Table 35. Multiplicities of irreducible representations for the 23AB-twisted sectors.

References

[1] J.H. Conway and S.P. Norton, "Monstrous Moonshine," Bull. London Math. Soc. **11**, 308 (1979).

[2] I. Frenkel, J. Lepowsky and A. Meurman, "Vertex operator algebras and the Monster," Pure and Applied Mathematics **134**, Academic Press, Inc., Boston, MA, (1988).

[3] R.E. Borcherds, "Monstrous Moonshine and monstrous Lie superalgebras," Invent. Math. **109**, 405 (1992).

[4] T.J. Gannon, "Moonshine beyond the Monster: The bridge connecting algebra, modular forms and physics," Cambridge University Presse (2006).

[5] T. Eguchi, H. Ooguri and Y. Tachikawa, "Notes on the K3 Surface and the Mathieu group M_{24}," Exper. Math. **20**, 91 (2011) [`arXiv:1004.0956` `[hep-th]`].

[6] M.C.N. Cheng, "K3 Surfaces, N=4 dyons, and the Mathieu group M_{24}," Commun. Number Theory Phys. **4**, 623 (2010) [`arXiv:1005.5415` `[hep-th]`].

[7] M.R. Gaberdiel, S. Hohenegger and R. Volpato, "Mathieu twining characters for K3," JHEP **1009**, 058 (2010) [`arXiv:1006.0221` `[hep-th]`].

[8] M.R. Gaberdiel, S. Hohenegger and R. Volpato, "Mathieu Moonshine in the elliptic genus of K3," JHEP **1010**, 062 (2010) [`arXiv:1008.3778` `[hep-th]`].

[9] T. Eguchi and K. Hikami, "Note on twisted elliptic genus of K3 surface," Phys. Lett. B **694**, 446 (2011) [`arXiv:1008.4924` `[hep-th]`].

[10] T. Gannon, "Much ado about Mathieu," `arXiv:1211.5531` `[math.RT]`.

[11] T. Creutzig, G. Höhn and T. Miezaki, "The McKay-Thompson series of Mathieu Moonshine modulo two," `arXiv:1211.3703v1` `[math.NT]`.

[12] M.C.N. Cheng and J.F.R. Duncan, "On Rademacher sums, the largest Mathieu group, and the holographic modularity of Moonshine," `arXiv:1110.3859` `[math.RT]`.

[13] M. R. Gaberdiel, S. Hohenegger and R. Volpato, "Symmetries of K3 sigma models," Comm. Number Theory and Phys. **6** (1) (2012) [`arXiv:1106.4315` `[hep-th]`].

[14] S. Norton, "Generalized moonshine," Proc. Sympos. Pure Math. **47**, 209 Amer. Math. Soc., Providence, RI (1987).

[15] C.-y. Dong, H.-s. Li and G. Mason, "Modular invariance of trace functions in orbifold theory," Commun. Math. Phys. **214**, 1 (2000) [`arXiv:q-alg/9703016`].

[16] G. Höhn, "Generalized moonshine for the baby Monster," Habilitation thesis, Mathematisches Institut, Universität Freiburg (2003), `http://www.math.ksu.edu/~gerald/papers/baby8.ps`.

[17] M.P. Tuite, "Generalized moonshine and abelian orbifold constructions," , Contemp. Math. **193**, 353 (1995) [`arXiv:hep-th/9412036`].

[18] R.I. Ivanov and M.P. Tuite, "Rational generalized moonshine from abelian orbifoldings of the moonshine module," Nucl. Phys. B **635**, 435 (2002) [`math/0106027` `[math-qa]`].

[19] R.I. Ivanov and M.P. Tuite, "Some irrational generalised moonshine from orbifolds," Nucl. Phys. B **635**, 473 (2002) [`math/0202275` `[math-qa]`].

[20] S. Carnahan, "Generalized Moonshine I: Genus zero functions," `arXiv:0812.3440` `[math.RT]`.

[21] S. Carnahan, "Generalized Moonshine II: Borcherds products," `arXiv:0908.4223` `[math.RT]`.

[22] S. Carnahan, "Generalized Moonshine III: Equivariant intertwining operators," to appear.

[23] S. Carnahan, "Generalized Moonshine IV: Monstrous Lie algebras," `arXiv:1208.6254 [math.RT]`.

[24] L.J. Dixon, P.H. Ginsparg and J.A. Harvey, "Beauty and the beast: superconformal symmetry in a Monster module," Commun. Math. Phys. **119**, 221 (1988).

[25] R. Dijkgraaf and E. Witten, "Topological gauge theories and group cohomology," Commun. Math. Phys. **129**, 393 (1990).

[26] T. Gannon, talk at 'Mathieu Moonshine', ETH Zürich, July 2011.

[27] G. Mason, "Orbifold conformal field theory and the cohomology of the Monster," unpublished.

[28] G. Mason, "On a system of elliptic modular forms attached to the largest Mathieu group," Nagoya Math. J. **118**, 177 (1990).

[29] D. Persson and R. Volpato, "Second Quantized Mathieu Moonshine," arXiv:1312.0622 [hep-th].

[30] M. Eichler and D. Zagier, "The theory of Jacobi forms," Birkhäuser (1985).

[31] T. Eguchi and A. Taormina, "Character formulas for the N=4 superconformal algebra," Phys. Lett. B **200**, 315 (1988).

[32] M. Dutour Sikirić and G. Ellis, "Wythoff polytopes and low-dimensional homology of Mathieu groups," J. Algebra **322**, 4143 (2009).

[33] C. Vafa, "Modular invariance and discrete torsion on orbifolds," Nucl. Phys. B **273**, 592 (1986).

[34] S. Willerton, "The twisted Drinfeld double of a finite group via gerbes and finite groupoids," Algebr. Geom. Topol. **8**, 1419 (2008) `[arXiv:math/0503266]`.

[35] P. Bantay, "Orbifolds and Hopf algebras," Phys. Lett. B **245**, 477 (1990).

[36] R. Dijkgraaf, V. Pasquier and P. Roche, "Quasi-Hopf algebras, group cohomology and orbifold models," Nucl. Phys. Proc. Suppl. **18B**, 60 (1990).

[37] A. Coste, T. Gannon and P. Ruelle, "Finite group modular data," Nucl. Phys. B **581**, 679 (2000) `[arXiv:hep-th/0001158]`.

[38] J. Fröhlich, J. Fuchs, I. Runkel and C. Schweigert, "Defect lines, dualities, and generalised orbifolds," in *XVIth International Congress on Mathematical Physics*, 608–613, World Sci. Publ., Hackensack, NJ (2010) `[arXiv:0909.5013 [math-ph]]`.

[39] K.S. Narain, M.H. Sarmadi and C. Vafa, "Asymmetric orbifolds," Nucl. Phys. B **288**, 551 (1987).

[40] M.R. Gaberdiel and R. Volpato, "Mathieu moonshine and orbifold K3s," `arXiv:1206.5143 [hep-th]`.

[41] The GAP Group, "GAP – Groups, Algorithms, and Programming, version 4.5.6," (2012) `http://www.gap-system.org`.

[42] G. Ellis, "HAP (Homological Algebra Programming) – a GAP package, version 1.10.8," (2012) `http://www.gap-system.org/Packages/hap.html`.

[43] K.S. Brown, "Cohomology of groups," Graduate Texts in Mathematics **87**, Springer, New York (1994).

[44] M.D.F. de Wild Propitius, "Topological interactions in broken gauge theories," `hep-th/9511195`.

[45] See the ancillary files attached with the preprint version of this paper at `http://arxiv.org`.

[46] S. Mukai, "Finite groups of automorphisms of $K3$ surfaces and the Mathieu group," Invent. Math. **94**, 183 (1988).

[47] A. Sen, "Discrete information from CHL black holes," JHEP **1011**, 138 (2010) `[arXiv:1002.3857 [hep-th]]`.

[48] R. Dijkgraaf, E.P. Verlinde and H.L. Verlinde, "Counting dyons in N=4 string theory," Nucl. Phys. B **484**, 543 (1997) [`arXiv:hep-th/9607026`].

[49] D. Shih, A. Strominger and X. Yin, "Recounting dyons in N=4 string theory," JHEP **0610**, 087 (2006) [`arXiv:hep-th/0505094`].

[50] A. Dabholkar, S. Murthy and D. Zagier, "Quantum black holes, wall crossing, and mock modular forms," `arXiv:1208.4074` [`hep-th`].

[51] S. Govindarajan, "BKM Lie superalgebras from counting twisted CHL dyons," JHEP **1105**, 089 (2011) [`arXiv:1006.3472` [`hep-th`]].

[52] S. Govindarajan, "Brewing moonshine for Mathieu," `arXiv:1012.5732` [`math.NT`].

[53] S. Govindarajan, "Unravelling Mathieu moonshine," Nucl. Phys. B **864**, 823 (2012) [`arXiv:1106.5715` [`hep-th`]].

[54] J.F. Duncan and I.B. Frenkel, "Rademacher sums, moonshine and gravity," Commun. Num. Theor. Phys. **5**, 849 (2011) [`arXiv:0907.4529` [`math.RT`]].

[55] V.A. Gritsenko and V.V. Nikulin, "Siegel automorphic form corrections of some Lorentzian Kac–Moody Lie algebras," [`arXiv:alg-geom/9504006`].

[56] V.A. Gritsenko and F. Clery, "The Siegel modular forms of genus 2 with the simplest divisor," [`arXiv:0812.3962` [`math.NT`]].

[57] M.C.N. Cheng and E.P. Verlinde, "Wall crossing, discrete attractor flow, and Borcherds algebra," SIGMA **4**, 068 (2008) [`arXiv:0806.2337` [`hep-th`]].

[58] M.C.N. Cheng and A. Dabholkar, "Borcherds-Kac-Moody symmetry of N=4 dyons," Commun. Num. Theor. Phys. **3**, 59 (2009) [`arXiv:0809.4258` [`hep-th`]].

[59] S. Govindarajan and K. Gopala Krishna, "Generalized Kac-Moody algebras from CHL dyons," JHEP **0904**, 032 (2009) [`arXiv:0807.4451` [`hep-th`]].

[60] S. Govindarajan and K. Gopala Krishna, "BKM Lie superalgebras from dyon spectra in Z_N CHL orbifolds for composite N," JHEP **1005**, 014 (2010) [`arXiv:0907.1410` [`hep-th`]].

[61] M.C.N. Cheng, J.F.R. Duncan and J.A. Harvey, "Umbral moonshine," `arXiv:1204.2779` [`math.RT`].

[62] J.A. Harvey and G.W. Moore, "Algebras, BPS states, and strings," Nucl. Phys. B **463**, 315 (1996) [`arXiv:hep-th/9510182`].

[63] J.A. Harvey and G.W. Moore, "On the algebras of BPS states," Commun. Math. Phys. **197**, 489 (1998) [`arXiv:hep-th/9609017`].

[64] A. Kapustin, "Chiral de Rham complex and the half-twisted sigma-model," `arXiv:hep-th/0504074`.

[65] F. Malikov, V. Schechtman and A. Vaintrob, "Chiral de Rham complex," Commun. Math. Phys. **204**, 439 (1999) [`arXiv:math/9803041`].

[66] L.A. Borisov and L. Libgober, "Elliptic genera of toric varieties and applications to mirror symmetry," Invent. Math. **140**, 453 (2000) [`arXiv:math/9904126`].

[67] G. Karpilovsky, "Group representations: volume 2," North-Holland Mathematical Studies **177**, Elsevier (1993).

Institut für Theoretische Physik, ETH Zürich, CH-8093 Zürich, Switzerland
E-mail address: gaberdiel@itp.phys.ethz.ch

Fundamental Physics, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden
E-mail address: daniel.persson@chalmers.se
URL: http://www.danper.se

Max-Planck-Institut für Dynamik und Selbstorganisation (MPIDS), 37077 Göttingen, Germany
E-mail address: henrik.ronellenfitsch@ds.mpg.de

Max-Planck-Institut für Gravitationsphysik, Am Mühlenberg 1, 14476 Golm, Germany
E-mail address: roberto.volpato@aei.mpg.de

arXiv:1006.0221v1 [hep-th] 1 Jun 2010

Preprint typeset in JHEP style - HYPER VERSION

Mathieu twining characters for K3

Matthias R. Gaberdiel, Stefan Hohenegger and Roberto Volpato

Institut für Theoretische Physik
ETH Zurich
CH-8093 Zürich
Switzerland

Abstract: The analogue of the McKay-Thompson series for the proposed Mathieu group action on the elliptic genus of K3 is analysed. The corresponding NS-sector twining characters have good modular properties and satisfy remarkable replication identities. These observations provide strong support for the conjecture that the elliptic genus of K3 carries indeed an action of the Mathieu group $\mathbb{M}_{24}$.

Contents

1. **Introduction and summary** **1**

2. **Elliptic genera and their modular properties** **2**
   2.1 Symmetric powers 4

3. **Mathieu symmetry and twining characters** **6**
   3.1 Explicit formulae for the twining characters 7
   3.2 Replication formulae 10

4. **Conclusions** **13**

A. **Definitions** **14**
   A.1 McKay-Thompson series 14
   A.2 Twining characters for K3 15

1. Introduction and summary

Recently, Eguchi, Ooguri and Tachikawa [11] have observed that the elliptic genus of K3 seems to carry an action of the Mathieu group $\mathbb{M}_{24}$. More specifically, they noticed that the expansion coefficients of the elliptic genus of K3, when written in terms of the elliptic genera of $\mathcal{N} = 4$ superconformal representations, can be expressed in terms of dimensions of Mathieu group representations. This intriguing observation is very reminiscent of the famous observation of McKay and Thompson who found that the Fourier expansion coefficients of the J-function can be written in terms of dimensions of representations of the Monster group. This led to a development that is now usually referred to as 'Monstrous Moonshine', see [18] for a nice review. One of the upshots of this analysis is that the J-function can be thought of as the partition function of a self-dual conformal field theory, the 'Monster conformal field theory' [3, 16], whose automorphism group is precisely the Monster group.

Unlike the situation of the Monster group, the dimensions of the Mathieu group are rather small, and thus there is a certain amount of arbitrariness in the above observation regarding the elliptic genus of K3. It would therefore be very interesting to subject this proposal to some more stringent tests. In the context of Monstrous Moonshine, one of the tests concerned the so-called McKay-Thompson series [21]. They are obtained from the J-function by replacing the dimensions of the Monster group representations by their characters with respect to a Monster group element. (In retrospect, they are the 'twining characters', where one inserts into the partition function of the Monster conformal field theory a Monster group element.) It was observed early on by Conway & Norton that these McKay-Thompson series have nice modular properties and satisfy remarkable replication identities [5]. The former identities have a natural interpretation in terms of standard orbifold arguments [22], while the latter can be obtained by considering the twining characters for the symmetric power theories (as follows from our analysis in section 3.2).

In this paper we study the analogues of the McKay-Thompson series for the case of K3, *i.e.* the twining elliptic genera where we insert a Mathieu group element into the trace. Using the proposed decomposition of the coefficients of the elliptic genus in terms of Mathieu group representations, we can calculate the first few coefficients of these twining genera explicitly. As for the McKay-Thompson series, standard orbifold arguments suggest that they must have good modular properties. At least for the group elements of small order, these modular constraints are very constraining, and we can use them to determine the twining genera exactly. We furthermore show that they satisfy highly non-trivial replication identities, relating twining genera corresponding to different group elements to one another. The fact that all of this works out represents a very stringent consistency check for the proposal of [11]; indeed, our analysis implies that one of the identifications of the coefficients in terms of Mathieu representation dimensions must be different to what was proposed in [11]. Our analysis thus gives strong support to the idea that the elliptic genus of K3 carries indeed an action of the Mathieu group $\mathbb{M}_{24}$. It furthermore suggests that this is in some sense the natural supersymmetric generalisation of Monstrous Moonshine to the case of $\mathbb{M}_{24}$. (A different proposal for the analogue of Monstrous Moonshine for $\mathbb{M}_{24}$ was made some time ago in the context of self-dual bosonic conformal field theories by Dong & Mason [8].)

The paper is organised as follows. In section 2 we recall some basic properties of elliptic genera and their modular properties. We also review how the elliptic genus of the symmetric power theory can be calculated in terms of the original elliptic genus. In section 3 we briefly describe the proposal of [11], and then calculate the twining elliptic genera and their associated NS twining characters. We explain in detail which modular properties the latter should possess. Using this insight we then identify the twining characters for group elements of small order in terms of known theta functions (or standard McKay-Thompson series), see (3.14). In section 3.2 we explain what replication formulae one should expect, and demonstrate in a large number of examples, that these identities are indeed satisfied for the twining characters of K3. Our conclusions and avenues for future research are described in section 4. We have relegated some of the more technical material to an appendix.

2. Elliptic genera and their modular properties

Let us begin by reviewing some standard material about elliptic genera and their Jacobi and modular properties. The elliptic genus of an $\mathcal{N}=2$ superconformal algebra is defined by

$$\phi(\tau,z) = \mathrm{Tr}_{\mathcal{H}_{RR}}\Big(q^{L_0-\frac{c}{24}} e^{2\pi i z J_0} (-1)^F \, \bar{q}^{\bar{L}_0-\frac{\bar{c}}{24}} (-1)^{\bar{F}}\Big) \ . \tag{2.1}$$

For the right-movers (whose modes are denoted by a bar), only the ground states contribute, and hence the above expression is in fact independent of $\bar{q}$. As is well-known [20], the modularity properties of conformal field theory together with spectral flow invariance and unitarity imply that the elliptic genus is a *weak Jacobi form* of index $m = \frac{c}{6}$ and weight zero [14]. A weak Jacobi form $\phi(\tau, z)$ of weight w and index $m \in \mathbb{Z}$, with $(\tau, z) \in \mathbb{H} \times \mathbb{C}$, satisfies the transformation laws

$$\phi\Big(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d}\Big) = (c\tau + d)^w \, e^{2\pi i m \frac{cz^2}{c\tau+d}} \, \phi(\tau, z) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) \ , \tag{2.2}$$

$$\phi(\tau, z + \ell\tau + \ell') = e^{-2\pi i m(\ell^2\tau + 2\ell z)} \phi(\tau, z) \qquad \ell, \ell' \in \mathbb{Z} \ , \tag{2.3}$$

and has a Fourier expansion

$$\phi(\tau, z) = \sum_{n \geq 0, \ell \in \mathbb{Z}} c(n, \ell) q^n y^\ell \tag{2.4}$$

with $c(n, \ell) = (-1)^w c(n, -\ell)$. For the case of K3 that will concern us primarily in this paper, the elliptic genus equals [10]

$$\phi_{K3}(\tau, z) = 2y + 20 + 2y^{-1} + q\Big(20y^2 - 128y + 216 - 128y^{-1} + 20y^{-2}\Big) + \mathcal{O}(q^2) \ , \tag{2.5}$$

where $y = e^{2\pi i z}$. For the following it will be convenient to rescale the elliptic genus by a factor of $\frac{1}{2}$, *i.e.* to define

$$\phi_{1\mathrm{A}}(\tau, z) = \frac{1}{2}\phi_{K3}(\tau, z) = y + 10 + y^{-1} + q\Big(10y^2 - 64y + 108 - 64y^{-1} + 10y^{-2}\Big) + \mathcal{O}(q^2) \ , \tag{2.6}$$

which is then directly equal to the standard weak Jacobi form $\phi_{0,1}$.

To understand the analogy with the bosonic Monster conformal field theory, it is useful to consider instead the NS-sector version of the elliptic genus, which is obtained by applying one half unit of spectral flow to the left-movers, *i.e.* $l = l' = \frac{1}{2}$ in (2.3); the shift by $l = \frac{1}{2}$ is familiar from the spectral flow automorphism

$$L_0 \mapsto L_0 + l\, J_0 + l^2 \, \frac{c}{6} \ , \qquad J_0 \mapsto J_0 + l\, \frac{c}{3} \ , \tag{2.7}$$

while the shift by $l' = \frac{1}{2}$ is required so as to remove the $(-1)^F$ from the trace, in order to get the standard NS-sector partition function without an insertion of $(-1)^F$. The resulting NS-sector partition function will always be denoted by

$$\chi(\tau, z) = \exp\left[2\pi i m \left(\frac{\tau}{4} + z + \frac{1}{2}\right)\right] \phi\left(\tau, z + \frac{\tau}{2} + \frac{1}{2}\right) \ . \tag{2.8}$$

Using the transformation properties of a Jacobi form it follows easily that

$$\begin{aligned} \chi(-1/\tau, z/\tau) &= (-1)^m e^{2\pi i \frac{mz^2}{\tau}} \, \chi(\tau, z) \\ \chi(\tau + 2, z) &= (-1)^m \, \chi(\tau, z) \ . \end{aligned} \tag{2.9}$$

In order to obtain a standard modular form (rather than a Jacobi form) we may then put $z = 0$; the resulting function (which we shall simply denote by $\chi(\tau)$) is non-trivial, and it has simple transformation laws under the congruence subgroup $\Gamma_\theta = \langle T^2, S\rangle$. For m even we have a strict modular function and for m odd we have a function with multiplier system given by -1 on the two generators [17],

$$\chi(-1/\tau) = (-1)^m \chi(\tau) \ , \qquad \chi(\tau+2) = (-1)^m \chi(\tau) \ . \tag{2.10}$$

For the case of K3 it is explicitly given as [17]

$$\chi_{1\mathrm{A}}(\tau) = \frac{1}{2}\chi_{K3}(\tau) = \kappa(\tau) = \left(\frac{2\vartheta_4(\tau)}{\vartheta_2(\tau)}\right)^2 - \left(\frac{2\vartheta_2(\tau)}{\vartheta_4(\tau)}\right)^2 = q^{-\frac{1}{4}}(1 - 20\, q^{\frac{1}{2}} + \cdots) \ . \tag{2.11}$$

Our conventions regarding the theta functions are described in the appendix.

It should be noted that specialising to $z = 0$ in the NS-sector does not loose any information. Indeed, the NS character $\chi_{1\mathrm{A}}(\tau)$ determines $\phi_{1\mathrm{A}}(\tau, \frac{1}{2})$ by setting $z = 0$ in

$$\chi_{1\mathrm{A}}(\tau, z) = \phi_{1\mathrm{A}}(\tau, 0)\, \frac{\vartheta_4(\tau, z)^2}{\vartheta_2(\tau, 0)^2} - \phi_{1\mathrm{A}}(\tau, \tfrac{1}{2}) \frac{\vartheta_3(\tau, z)^2}{\vartheta_2(\tau, 0)^2} \ , \tag{2.12}$$

and using that the coefficient of the first term is simply the Witten index $\phi_{1\mathrm{A}}(\tau, 0) = 12$. On the other hand, this information determines the full elliptic genus via

$$\phi_{1\mathrm{A}}(\tau, z) = \phi_{1\mathrm{A}}(\tau, 0)\, \frac{\vartheta_2(\tau, z)^2}{\vartheta_2(\tau, 0)^2} + \phi_{1\mathrm{A}}(\tau, \tfrac{1}{2})\, \frac{\vartheta_1(\tau, z)^2}{\vartheta_2(\tau, 0)^2} \ . \tag{2.13}$$

Thus the knowledge of $\chi_{1\mathrm{A}}(\tau)$ is equivalent to that of $\phi_{1\mathrm{A}}(\tau, z)$. The advantage of working with $\chi_{1\mathrm{A}}(\tau)$ is that it is the natural supersymmetric analogue of the partition function of a bosonic conformal field theory for which Monstrous Moonshine was originally formulated.

2.1 Symmetric powers

For the following it will also be important to consider the symmetric powers of the K3 conformal field theory. Given any $\mathcal{N} = 2$ superconformal field theory $\mathcal{H}$, we can consider the symmetric power theory

$$\mathcal{H}^{(p)} = \underbrace{\Big(\mathcal{H} \otimes \cdots \otimes \mathcal{H}\Big)}_{\text{p times}} / \mathbb{Z}_p \ , \tag{2.14}$$

where the generator π of $\mathbb{Z}_p$ is the cyclic permutation of length p. We shall concentrate in the following on the case when p is prime; then all elements of $\mathbb{Z}_p$ are cyclic permutations of cycle length p. Given the elliptic genus of the original $\mathcal{H}$ theory, we can then determine the elliptic genus of $\mathcal{H}^{(p)}$ [6]. From the untwisted sector of the $\mathbb{Z}_p$ orbifold we get the contribution

$$\phi_U^{(p)}(\tau, z) = \frac{1}{p}\Big(\big(\phi(\tau, z)\big)^p + (p-1)\, \phi(p\tau, pz) \Big) \ . \tag{2.15}$$

Here the first term is the elliptic genus without the insertion of any orbifold generator, while the second term gives the contributions coming from π^l with $l = 1, \ldots, p-1$. Since each such π^l is a cyclic permutation of order p, only the states of the form $(u \otimes u \otimes \cdots \otimes u)$ contribute, and this leads to $\phi(p\tau, pz)$ for each of these terms.

There are $(p-1)$ twisted sectors, and each of them contributes the same, thus giving for the whole twisted sector

$$\phi_T^{(p)}(\tau, z) = \frac{(p-1)}{p} \sum_{l=0}^{p-1} \phi\big(\tfrac{\tau+l}{p}, z\big) \ , \tag{2.16}$$

where the sum over l implements the orbifold projection in the twisted sector. Thus the elliptic genus of the symmetric power theory is (for p prime)

$$\phi^{(p)}(\tau, z) = \frac{1}{p}\,\big(\phi(\tau, z)\big)^p + \frac{(p-1)}{p}\, H_p\phi(\tau, z) \ , \tag{2.17}$$

where H_p denotes the action of the Hecke operator

$$H_p\phi(\tau, z) = \phi(p\tau, pz) + \sum_{l=0}^{p-1} \phi\big(\tfrac{\tau+l}{p}, z\big) \ . \tag{2.18}$$

It is well known that (2.17) is again a weak Jacobi form of index mp provided that $\phi(\tau, z)$ has index m [6]. For example, this follows directly from the fact that the Hecke operator (2.18) maps weak Jacobi forms to weak Jacobi forms, see [14].

For our purposes, it is also convenient to rewrite (2.18) for the NS-sector character $\chi^{(p)}(\tau)$. The result depends on whether $p = 2$ or p odd. For $p = 2$ we find

$$\chi^{(2)}(\tau) = \chi\left(2\tau, \tfrac{1}{2}\right) + \phi\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) + (-1)^m \phi\left(\tfrac{\tau}{2} + \tfrac{1}{2}, 0\right) \ , \tag{2.19}$$

where $\chi(\tau, z)$ is the NS-sector function defined in (2.8), while $\phi(\tau, z)$ is the elliptic genus in the R-sector. In particular, the last expression is therefore actually a constant since $\phi(\tau, 0)$ is the Witten index. The emergence of the R-sector as the 'twisted' sector of the NS-sector is familiar from standard $\mathbb{Z}_2$ orbifold considerations. For the case of the K3 character the relevant formula is[1]

$$\begin{aligned} \chi_{1\mathrm{A}}^{(2)}(\tau) &= \left[\left(\frac{2\vartheta_3(2\tau)}{\vartheta_2(2\tau)}\right)^2 - \left(\frac{2\vartheta_2(2\tau)}{\vartheta_3(2\tau)}\right)^2\right] + \left[\left(\frac{2\vartheta_3(\frac{\tau}{2})}{\vartheta_4(\frac{\tau}{2})}\right)^2 - \left(\frac{2\vartheta_4(\frac{\tau}{2})}{\vartheta_3(\frac{\tau}{2})}\right)^2\right] - 12 \\ &= \chi_{1\mathrm{A}}(\tau)^2 + 36 \ . \end{aligned} \tag{2.20}$$

The fact that the left-hand side can be expressed in terms of a polynomial of the original K3 character is a consequence of $\chi_{1\mathrm{A}}(\tau)^2$ being a Hauptmodul (without any multiplier system) for the congruence subgroup Γ_θ. Since the left-hand is a modular form for Γ_θ (without any multiplier system), it must therefore be a polynomial of $\chi_1(\tau)^2$.

For $p > 2$ (recall that p is prime, and therefore now odd) the formula is instead

$$\chi^{(p)}(\tau) = \chi(p\tau) + \sum_{l=0}^{p-1} \chi\big(\tfrac{\tau+4l}{p}\big) \ , \qquad p \geq 3 \text{ prime.} \tag{2.21}$$

---

[1] We shall state these and any following formulae for the rescaled elliptic genera, see (2.6) and (2.11).

For the case of the K3 character we then find, for example for $p = 3$,

$$\begin{aligned} \chi^{(3)}_{1\mathrm{A}}(\tau) &= \chi_{1\mathrm{A}}(3\tau) + \chi_{1\mathrm{A}}\left(\tfrac{\tau}{3}\right) + \chi_{1\mathrm{A}}\left(\tfrac{\tau+4}{3}\right) + \chi_{1\mathrm{A}}\left(\tfrac{\tau+8}{3}\right) \\ &= \chi_{1\mathrm{A}}(\tau)^3 + 60\,\chi_{1\mathrm{A}}(\tau)\ . \end{aligned} \tag{2.22}$$

By the same token as above, this formula follows from the fact that the left-hand side is again a modular form (with multiplier system -1 on S and T^2) for Γ_θ, and that the Hauptmodul for these is $\chi_{1\mathrm{A}}(\tau)$. Because of the non-trivial multipler system of the left-hand-side the expression now only involves odd powers of $\chi_{1\mathrm{A}}(\tau)$.

We should note that these relations are the natural analogue of the replication formulae for the J-function, which read for $p = 2$ and $p = 3$

$$\begin{aligned} J(2\tau) + J\left(\tfrac{\tau}{2}\right) + J\left(\tfrac{\tau+1}{2}\right) &= J^2(\tau) - 2a_1 \\ J(3\tau) + J\left(\tfrac{\tau}{3}\right) + J\left(\tfrac{\tau+1}{3}\right) + J\left(\tfrac{\tau+2}{3}\right) &= J^3(\tau) - 3a_1 J(\tau) - 3a_2\ . \end{aligned} \tag{2.23}$$

Here $J(\tau)$ is the famous J-function

$$J(\tau) = \sum_{k=-1}^{\infty} a_k q^k = q^{-1} + 196884q + 21493760q^2 + \mathcal{O}(q^3)\ , \tag{2.24}$$

that plays a central role in Monstrous moonshine, see [18] for a nice review.

3. Mathieu symmetry and twining characters

The elliptic genus of K3 can be thought of as the partition function of the $\mathcal{N} = 2$ half-BPS states of type II string theory on K3. It is suggested by the work of [11] that the space of these half-BPS states carries an action of the Mathieu group $\mathbb{M}_{24}$. Indeed, it was found in [11], following on from earlier work [9], that the elliptic genus of K3 can be written as

$$\phi_{1\mathrm{A}}(\tau, z) = 24\,\mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=0}(\tau, z) + \sum_{n=0}^{\infty} A_n\,\mathrm{ch}^{\mathcal{N}=4}_{h=n+\frac{1}{4},l=\frac{1}{2}}(\tau, z)\ , \tag{3.1}$$

where $\mathrm{ch}^{\mathcal{N}=4}_{h=\frac{1}{4},l=0}$ is the elliptic genus of the short $\mathcal{N} = 4$ representation with $h = \frac{1}{4}$ and $l = 0$ — see [12, 13] for an explicit formula — while

$$\mathrm{ch}^{\mathcal{N}=4}_{h,l=\frac{1}{2}}(\tau, z) = q^{h-\frac{3}{8}}\,\frac{\vartheta_1(\tau, z)^2}{\eta(\tau)^3} \tag{3.2}$$

is the elliptic genus of a long $\mathcal{N} = 4$ representation.[2] The observation of [11] was that the coefficients A_n can be written in terms of dimensions of irreducible representations of the Mathieu group $\mathbb{M}_{24}$. (The character table of $\mathbb{M}_{24}$, from which one can read off in particular the dimensions of the irreducible representations is given in the appendix.) More explicitly,

$$\begin{array}{rclcrcl} 24 &=& \mathbf{23} + \mathbf{1} & \qquad & -A_0 = 2 &=& \mathbf{1} + \mathbf{1} \\ A_1 = 90 &=& \mathbf{45} + \overline{\mathbf{45}} & & A_2 = 462 &=& \mathbf{231} + \overline{\mathbf{231}} \\ A_3 = 1440 &=& \mathbf{770} + \overline{\mathbf{770}} & & A_4 = 4554 &=& \mathbf{2277} + \overline{\mathbf{2277}} \\ A_5 = 11592 &=& 2 \cdot \mathbf{5796} & & A_6 = 27830 &=& 2 \cdot \mathbf{3520} + 2 \cdot \mathbf{10395} \end{array} \tag{3.3}$$

---

[2] Strictly speaking, the $\mathcal{N} = 4$ representation with $n = 0$ ($h = \frac{1}{4}$) is short, and thus (3.2) for $h = \frac{1}{4}$ is not the elliptic genus of a single representation, but rather involves a sum of representations.

and

$$A_7 = 61686 = 2 \cdot \mathbf{1771} + 2 \cdot \mathbf{2024} + 2 \cdot \mathbf{5313} + 2 \cdot \mathbf{5796} + 2 \cdot \mathbf{5544} + 2 \cdot \mathbf{10395} \ . \tag{3.4}$$

Note that, relative to [11], we have absorbed the prefactor of 2 in [11, eq. (1.11)] into the definition of A_n. Then we can write the A_n in terms of real representations, *i.e.* sums of representations and their conjugate representations. Furthermore, for reasons that will be explained momentarily, our expression for A_7 differs from the decomposition suggested in [11]. We should mention though that while A_7 given in (3.4) is consistent with what is done below, it is not uniquely specified by this.

3.1 Explicit formulae for the twining characters

Obviously, as is apparent from the last comment, there is much arbitrariness in these decompositions, and they are only a suggestive hint of some underlying deeper structure. The main aim of this paper is to subject this proposal to a much more refined test. To this end we shall study the corresponding '*twining genus*', *i.e.* the analogues of the McKay-Thompson series in the case of Monstrous Moonshine. This is to say, we shall calculate not just the elliptic genus itself, but the elliptic genus with the insertion of a group element $g \in \mathbb{M}_{24}$

$$\phi_g(\tau, z) = \mathrm{Tr}_{\mathcal{H}_{RR}} \Big(g \, q^{L_0 - \frac{c}{24}} e^{2\pi i z J_0} (-1)^F \, \bar{q}^{\bar{L}_0 - \frac{\bar{c}}{24}} (-1)^{\bar{F}} \Big) \ . \tag{3.5}$$

Technically speaking, the calculation of ϕ_g differs from ϕ by replacing A_n in (3.1) by the characters of g in the representations described in (3.3). Obviously, the character of g only depends on its conjugacy class, and thus all the information we need is contained in the character table of $\mathbb{M}_{24}$ that is included in the appendix.

One immediate problem with this approach is that we only know how to write A_n in terms of dimensions of representations of $\mathbb{M}_{24}$ for the first few values of n. We can therefore only directly determine the low powers of q of the twining genera ϕ_g. If this was all there was to it, this would not be particularly interesting, and it would certainly not lead to any non-trivial consistency check of the proposal. The reason why interesting information can be obtained in this manner is that the twining genera must have fairly specific modular properties. One way to see this explicitly is to translate the R-sector series $\phi_g(\tau, z)$ into the NS-sector as in (2.8), and set $z = 0$. The resulting expression is the NS-sector twining character $\chi_g(\tau)$, *i.e.* the character in the NS-sector with the insertion of the Mathieu group element g. Note that by the same arguments as above we are not loosing any information in doing so, because the analogues of (2.12) and (2.13) are

$$\chi_g(\tau, z) = \phi_g(\tau, 0) \, \frac{\vartheta_4(\tau, z)^2}{\vartheta_2(\tau, 0)^2} - \phi_g(\tau, \tfrac{1}{2}) \frac{\vartheta_3(\tau, z)^2}{\vartheta_2(\tau, 0)^2} \ , \tag{3.6}$$

and

$$\phi_g(\tau, z) = \phi_g(\tau, 0) \, \frac{\vartheta_2(\tau, z)^2}{\vartheta_2(\tau, 0)^2} + \phi_g(\tau, \tfrac{1}{2}) \, \frac{\vartheta_1(\tau, z)^2}{\vartheta_2(\tau, 0)^2} \ , \tag{3.7}$$

where now the constant term equals

$$\phi_g(\tau, 0) = \tfrac{1}{2} \, \mathrm{Tr}_{\mathbf{23} \oplus \mathbf{1}}(g) \ . \tag{3.8}$$

Thus we can reconstruct the full twining genus $\phi_g(\tau,z)$ from the twining character $\chi_g(\tau)$, and vice versa. In terms of the usual box notation, $\chi_g(\tau)$ corresponds to

$$\chi_g(\tau) \longleftrightarrow (-1,\mathbf{1})\ \underset{(-1,\mathbf{g})}{\boxed{}} \tag{3.9}$$

where the '-1's indicate that the trace is taken in the NS-sector, and without any insertion of $(-1)^F$, respectively. Under a modular transformation it is believed that these twining and twisted characters transform as

$$h\ \underset{g}{\boxed{}} \quad \xrightarrow{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \quad h^d g^c\ \underset{g^a h^b}{\boxed{}} \tag{3.10}$$

For the case at hand, it therefore follows that the twining character corresponding to (3.9) is invariant (possibly up to some phase) under a modular transformation provided that

$$a+b,\ c+d \in 2\mathbb{Z}+1\ , \qquad c = 0 \mod o(g), \tag{3.11}$$

where $o(g)$ is the order of the group element g. [Here we have used that the group operation is $(\epsilon_1, g_1)\cdot(\epsilon_2, g_2) = (\epsilon_1\epsilon_2, g_1\cdot g_2)$, where $\epsilon_j \in \{\pm 1\}$ and $g_i \in \mathbb{M}_{24}$. We have furthermore used that the conditions in (3.11) together with the requirement that $ad-bc=1$ imply that $(a,o(g))=1$. For the Mathieu group $\mathbb{M}_{24}$ it then follows that g^a is in the same conjugacy class as g.] The subgroup of $SL(2,\mathbb{Z})$ that is characterised by (3.11) contains in particular

$$\Gamma^g = \Gamma_\theta \cap \Gamma_0(o(g))\ , \tag{3.12}$$

where Γ_θ was introduced before, and

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})\ :\ c = 0 \bmod (N) \right\}\ . \tag{3.13}$$

The property of a function to be modular invariant under Γ^g (possibly up to some multiplier system) is a very strong condition, and knowing the first few terms usually determines the function uniquely.

Thus the non-trivial consistency test we will be performing proceeds as follows. Using the low-lying decomposition (3.3), we can calculate the first few coefficients of the twining characters $\chi_g(\tau)$ explicitly. On the other hand, we know that these functions must be invariant under Γ^g, and thus there is the highly non-trivial consistency condition whether this ansatz can be completed in a Γ^g-invariant manner. The results we find are consistent with A_n given in (3.3) and (3.4), but not with the formula for A_7 suggested in [11]. They represent a highly non-trivial consistency condition for the conjecture of [11] that the elliptic genus of K3 carries an action by the Mathieu group $\mathbb{M}_{24}$.

The constraints coming from the modular properties are tighter the larger Γ^g is, *i.e.* the smaller the order of the group element is. In particular, we should therefore expect to have good control over the twining characters corresponding to group elements of small order. Indeed, this turns out to be correct, and we have even found closed form expressions that match with the low lying coefficients and have the correct modular properties. More specifically, we find that — in the following the label refers to the conjugacy class of the Mathieu group element that is described in the appendix

$$\begin{aligned}
\chi_{1\mathrm{A}}(\tau) &= \kappa(\tau) = T_{[8\mathrm{a}]}(\tfrac{\tau}{4}) \qquad \text{— see eq. (2.11)} \\
\chi_{2\mathrm{A}}(\tau) &= 4\,\frac{\vartheta_4(\tau)^2}{\vartheta_2(\tau)^2} \simeq T_{[16\mathrm{b}]}(\tfrac{\tau}{4}) \\
\chi_{2\mathrm{B}}(\tau) &= 4\,\frac{\vartheta_3(\tau)^2}{\vartheta_2(\tau)^2} \simeq T_{[16\mathrm{b}]}(\tfrac{\tau}{4}) \\
\chi_{3\mathrm{A}}(\tau) &= \left(\frac{\eta(\tau)}{\eta(3\tau)}\right)^3 \frac{\vartheta_3(3\tau)}{\vartheta_3(\tau)} = T_{[24\mathrm{c}]}(\tfrac{\tau}{4}) \\
\chi_{4\mathrm{A}}(\tau) &= \left(\frac{\eta(\tau)\,\eta(2\tau)}{\eta(\frac{\tau}{2})\,\eta(4\tau)}\right)^4 = T_{[16\mathrm{A}]}(\tfrac{\tau}{4}) \\
\chi_{4\mathrm{B}}(\tau) &= 2\,\frac{\vartheta_3(2\tau)}{\vartheta_2(2\tau)} = T_{[16\mathrm{B}]}(\tfrac{\tau}{4}) \\
\chi_{4\mathrm{C}}(\tau) &= 4\,\frac{\vartheta_3(\tau)^2\vartheta_3(2\tau)}{\vartheta_2(\tau)^2\vartheta_4(2\tau)} \\
\chi_{5\mathrm{A}}(\tau) &= q^{-\frac{1}{4}}\Big(1+3\,q+4\,q^{\frac{3}{2}}+4\,q^2+4\,q^{\frac{5}{2}}+\cdots\Big) = T_{[40\mathrm{a}]}(\tfrac{\tau}{4}) \\
\chi_{6\mathrm{A}}(\tau) &= \left(\frac{\eta(\frac{3}{2}\tau)\,\eta(2\tau)}{\eta(\frac{\tau}{2})\,\eta(6\tau)}\right)^2 = T_{[24\mathrm{H}]}(\tfrac{\tau}{4})\,.
\end{aligned} \tag{3.14}$$

Our conventions for the Jacobi theta functions, and the first few coefficients of the power series expansions of the relevant McKay-Thompson series are given in the appendix. We have again normalised these functions so that they start with $1 \cdot q^{-\frac{1}{4}}$, compare (2.11). In most cases, the functions can be identified with a standard McKay-Thompson series $T_{[*]}$ (evaluated at $\frac{\tau}{4}$) — if this is the case the corresponding identity is given above. For the conjugacy class 2A and 2B the coefficients only agree with the McKay-Thompson series up to some signs; this is indicated by $\simeq$. In all cases we have explicitly checked that these functions are invariant under Γ^g in (3.12) up to some multiplier system. The multiplier system is non-trivial in all cases since all twining characters transform as -1 under $T^2 \in \Gamma^g$. In addition, $\chi_{1\mathrm{A}}$ transforms as -1 under S (compare (2.9)), while $\chi_{4\mathrm{C}}$ transforms as $-i$ under ST^4S. (We have not checked the precise form of the multiplier system for the class corresponding to 5A.)

The attentive reader will notice that (3.14) accounts for all twining characters of group elements up to order 6, except for the ones corresponding to 3B and 6B. While we have not managed to find any closed formulae for them — explicit formulae for the first few coefficients of their power series expansion are given in the appendix — we have an indirect argument that suggests that at least the twining character corresponding to 3B has good modular properties; this will be explained in the following section. We should also mention that some of these conjugacy classes are actually contained in $\mathbb{M}_{23}$ — these are the conjugacy classes 1A, 2A, 3A, 4B, 5A, 6A — while the others (2B, 3B, 4A, 4C) lie only in $\mathbb{M}_{24}$ but not in $\mathbb{M}_{23}$. These latter classes are characterised by the property that $\mathrm{Tr}_{\mathbf{23}\oplus\mathbf{1}}(g) = 0$. For them the relation between the twining characters and the twining genera (see (3.6) and (3.7)) is therefore particularly simple, and we have

$$\phi_g(\tau, z) = -\chi_g(\tau)\, \frac{\vartheta_1(\tau,z)^2}{\vartheta_3(\tau,0)^2} \qquad \text{if } \mathrm{Tr}_{\mathbf{23}\oplus\mathbf{1}}(g) = 0. \tag{3.15}$$

Thus for the elements in $\mathbb{M}_{24}$ that are not in $\mathbb{M}_{23}$, the modular properties of $\phi_g(\tau, z)$ are a direct consequence of those of $\chi_g(\tau)$. However, apart from this simplification, we do not see any difference between the behaviour of the elements that lie in $\mathbb{M}_{23}$ and those that do not.

3.2 Replication formulae

One of the remarkable properties of the McKay-Thompson series in Monstrous Moonshine are their replication relations. They are a direct generalisation of the J-function replication identities (2.23) to the McKay-Thompson series. In the following we want to give a conformal field theoretic argument that suggests that similar relations should also hold for the twining characters of K3, and more specifically, what precise form they should take. Incidentally, if we applied the following argument to the Monster theory it would give a physics explanation of the replication identities of Monstrous Moonshine.

The idea of the argument is very simple. If the $\mathcal{N} = 2$ half-BPS spectrum of K3 carries an action of the Mathieu group $\mathbb{M}_{24}$, then the same will apply to the symmetric power theories where $g \in \mathbb{M}_{24}$ acts diagonally on all p factors so as to commute with the orbifold action. But if this is so, then we can also consider the twining character corresponding to $g \in \mathbb{M}_{24}$ for the p^{th} symmetric power theory. Given our detailed understanding of the orbifold theory, we can determine how g acts on the various terms in (2.17) or indeed on the different terms in $H_p\phi(\tau, z)$. The first term corresponds to the untwisted sector contribution where the cyclic permutation generator has been inserted into the trace. This implies that only the states of the form $(u \otimes u \otimes \cdots \otimes u)$ contribute. Inserting the diagonal group element $g \in \mathbb{M}_{24}$ then leads to $\phi_{g^p}(p\tau, pz)$. The other terms of $H_p\phi(\tau, z)$ come from the twisted sector, and the natural action of $g \in \mathbb{M}_{24}$ is simply described by replacing ϕ by ϕ_g. Thus we are led to claim that the twining genus for the Hecke part of the p^{th} symmetric power theory is given by

$$\big(H_p\phi\big)_g(\tau, z) = \phi_{g^p}(p\tau, pz) + \sum_{l=0}^{p-1} \phi_g\big(\tfrac{\tau+l}{p}, z\big) \ . \tag{3.16}$$

Translated into the NS-sector characters we then have for $p = 2$

$$\chi^{(2)}_g(\tau) = \chi_{g^2}\left(2\tau, \tfrac{1}{2}\right) + \phi_g\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) + (-1)^m \phi_g\left(\tfrac{\tau}{2} + \tfrac{1}{2}, 0\right) \ , \tag{3.17}$$

while for $p > 2$ the formula is

$$\chi_g^{(p)}(\tau) = \chi_{g^p}(p\tau) + \sum_{l=0}^{p-1} \chi_g\left(\tfrac{\tau+4l}{p}\right) , \qquad p \geq 3 \text{ prime.} \tag{3.18}$$

On the other hand, it is clear from the discussion around (3.12) that also these twining characters must be invariant under Γ^g, possibly up to a (different) multiplier system. In the context of Monstrous Moonshine, the McKay-Thompson series were Hauptmoduls for the corresponding genus zero congruence subgroups, and thus the McKay-Thompson series of the symmetric power theories could be written as a polynomial of the original McKay-Thompson series. This then leads to the replication formulae of Monstrous Moonshine (generalising directly the J-function replication identities of (2.23)).

In the current context, it appears that the twining characters are *not* of genus zero, but we still find remarkable replication identities that give yet another consistency check on our analysis (and indicate that the twining characters corresponding to 3B and 7A=7B have indeed good modular properties). Specifically, we have found the following identities for $p = 2$

$$\chi_{2\mathrm{A}}^{(2)}(\tau) = \chi_{1\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{2\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - 4 = \chi_{2\mathrm{A}}(\tau)^2 + 4 \tag{3.19}$$

$$\chi_{2\mathrm{B}}^{(2)}(\tau) = \chi_{1\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{2\mathrm{B}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) = \chi_{2\mathrm{B}}(\tau)^2 - 12 \tag{3.20}$$

$$\chi_{3\mathrm{A}}^{(2)}(\tau) = \chi_{3\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{3\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - 3 = \chi_{3\mathrm{A}}(\tau)^2 \tag{3.21}$$

$$\chi_{3\mathrm{B}}^{(2)}(\tau) = \chi_{3\mathrm{B}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{3\mathrm{B}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) = -\tfrac{1}{2}\,\chi_{3\mathrm{B}}(\tau)^2 + \tfrac{3}{2}\,T_{[12\mathrm{D}]}\left(\tfrac{\tau}{2}\right) \tag{3.22}$$

$$\chi_{4\mathrm{A}}^{(2)}(\tau) = \chi_{2\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{4\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) = -\chi_{4\mathrm{A}}(\tau)^2 + 2\,T_{[8\mathrm{E}]}\left(\tfrac{\tau}{2}\right) + 4 \tag{3.23}$$

$$\chi_{4\mathrm{B}}^{(2)}(\tau) = \chi_{2\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{4\mathrm{B}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - 2 = -\chi_{4\mathrm{B}}(\tau)^2 + 2\,T_{[8\mathrm{E}]}\left(\tfrac{\tau}{2}\right) - 4 \tag{3.24}$$

$$\chi_{4\mathrm{C}}^{(2)}(\tau) = \chi_{2\mathrm{B}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{4\mathrm{C}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) = -\tfrac{1}{2}\,\chi_{4\mathrm{C}}(\tau)^2 + \tfrac{3}{2}\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right)^2 \tag{3.25}$$

$$\chi_{5\mathrm{A}}^{(2)}(\tau) = \chi_{5\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{5\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - 2 = -\chi_{5\mathrm{A}}(\tau)^2 + 2\,T_{[20\mathrm{C}]}\left(\tfrac{\tau}{2}\right) - 4 \tag{3.26}$$

$$\chi_{6\mathrm{A}}^{(2)}(\tau) = \chi_{3\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{6\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - 1 = -\chi_{6\mathrm{A}}(\tau)^2 + 2\,T_{[12\mathrm{I}]}\left(\tfrac{\tau}{2}\right) \tag{3.27}$$

$$\chi_{7\mathrm{A}}^{(2)}(\tau) = \chi_{7\mathrm{A}}\left(2\tau, \tfrac{1}{2}\right) + \phi_{7\mathrm{A}}\left(\tfrac{\tau}{2}, \tfrac{1}{2}\right) - \tfrac{3}{2} = -2\chi_{7\mathrm{A}}(\tau)^2 + 3\,T_{[28\mathrm{B}]}\left(\tfrac{\tau}{2}\right) . \tag{3.28}$$

These identities imply recursion relations for the twining characters which one could use, say, to determine the twining character for 3B explicitly. Furthermore, since the McKay-Thompson series are known to have good modular transformation properties, the existence of these relations implies that the same must be true for the twining characters that appear here.

For $p = 3$ we have found the following explicit identities

$$\chi_{2\mathrm{A}}^{(3)}(\tau) = \chi_{2\mathrm{A}}(3\tau) + \sum_{l=0}^{2} \chi_{2\mathrm{A}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{2\mathrm{A}}(\tau)^3 + 12\,\chi_{2\mathrm{A}}(\tau) \tag{3.29}$$

$$\chi^{(3)}_{\mathrm{2B}}(\tau) = \chi_{\mathrm{2B}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{2B}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{2B}}(\tau)^3 - 12\,\chi_{\mathrm{2B}}(\tau) \tag{3.30}$$

$$\chi^{(3)}_{\mathrm{3A}}(\tau) = \chi_{\mathrm{1A}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{3A}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{3A}}(\tau)^3 + 6\,\chi_{\mathrm{3A}}(\tau) \tag{3.31}$$

$$\chi^{(3)}_{\mathrm{4A}}(\tau) = \chi_{\mathrm{4A}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{4A}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{4A}}(\tau)^3 - 12\,\chi_{\mathrm{4A}}(\tau) \tag{3.32}$$

$$\chi^{(3)}_{\mathrm{4B}}(\tau) = \chi_{\mathrm{4B}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{4B}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{4B}}(\tau)^3 \tag{3.33}$$

$$\chi^{(3)}_{\mathrm{4C}}(\tau) = \chi_{\mathrm{4C}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{4C}}\left(\tfrac{\tau+4l}{3}\right) = \tfrac{1}{4}\chi_{\mathrm{4C}}(\tau)^3 + \tfrac{3}{4}\chi_{\mathrm{4C}}(\tau)\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right)^2 - 6\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right) \tag{3.34}$$

$$\chi^{(3)}_{\mathrm{5A}}(\tau) = \chi_{\mathrm{5A}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{5A}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{5A}}(\tau)^3 \tag{3.35}$$

$$\chi^{(3)}_{\mathrm{6A}}(\tau) = \chi_{\mathrm{2A}}(3\tau) + \sum_{l=0}^{2} \chi_{\mathrm{6A}}\left(\tfrac{\tau+4l}{3}\right) = \chi_{\mathrm{6A}}(\tau)^3 - 6\,\chi_{\mathrm{6A}}(\tau)\ , \tag{3.36}$$

while for $p = 5$ we have found

$$\chi^{(5)}_{\mathrm{2A}}(\tau) = \chi_{\mathrm{2A}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{2A}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{2A}}(\tau)^5 + 20\,\chi_{\mathrm{2A}}(\tau)^3 + 70\,\chi_{\mathrm{2A}}(\tau) \tag{3.37}$$

$$\chi^{(5)}_{\mathrm{2B}}(\tau) = \chi_{\mathrm{2B}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{2B}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{2B}}(\tau)^5 - 20\,\chi_{\mathrm{2B}}(\tau)^3 + 70\,\chi_{\mathrm{2B}}(\tau) \tag{3.38}$$

$$\chi^{(5)}_{\mathrm{3A}}(\tau) = \chi_{\mathrm{3A}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{3A}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{3A}}(\tau)^5 + 10\,\chi_{\mathrm{3A}}(\tau)^3 + 15\,\chi_{\mathrm{3A}}(\tau) \tag{3.39}$$

$$\chi^{(5)}_{\mathrm{4A}}(\tau) = \chi_{\mathrm{4A}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{4A}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{4A}}(\tau)^5 - 20\,\chi_{\mathrm{4A}}(\tau)^3 + 30\,\chi_{\mathrm{4A}}(\tau) \tag{3.40}$$

$$\chi^{(5)}_{\mathrm{4B}}(\tau) = \chi_{\mathrm{4B}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{4B}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{4B}}(\tau)^5 - 10\,\chi_{\mathrm{4B}}(\tau) \tag{3.41}$$

$$\begin{aligned}\chi^{(5)}_{\mathrm{4C}}(\tau) = \chi_{\mathrm{4C}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{4C}}\left(\tfrac{\tau+4l}{5}\right) &= \tfrac{1}{16}\,\chi_{\mathrm{4C}}(\tau)^5 - \tfrac{1}{16}\chi_{\mathrm{4C}}(\tau)^3\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right)^2 \\ &+\chi_{\mathrm{4C}}(\tau)\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right)^4 - \tfrac{1}{2}\,\chi_{\mathrm{4C}}(\tau)^2\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right) - 4\,T_{[16\mathrm{d}]}\left(\tfrac{\tau}{4}\right)^3 - 2\,\chi_{\mathrm{4C}}(\tau)\end{aligned} \tag{3.42}$$

$$\chi^{(5)}_{\mathrm{5A}}(\tau) = \chi_{\mathrm{1A}}(5\tau) + \sum_{l=0}^{4} \chi_{\mathrm{5A}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{\mathrm{5A}}(\tau)^5 - 15\,\chi_{\mathrm{5A}}(\tau) \tag{3.43}$$

$$\chi_{6\mathrm{A}}^{(5)}(\tau) = \chi_{6\mathrm{A}}(5\tau) + \sum_{l=0}^{4} \chi_{6\mathrm{A}}\left(\tfrac{\tau+4l}{5}\right) = \chi_{6\mathrm{A}}(\tau)^5 - 10\,\chi_{6\mathrm{A}}(\tau)^3 - 5\,\chi_{6\mathrm{A}}(\tau)\ . \tag{3.44}$$

These are probably not the only identities that exist, but only those we have been able to find with our current partial understanding of the twining characters. Indeed, in order to be able to determine the higher order replication formulae, one needs to know the twining characters up to some fairly high level. Thus with our current understanding we can only check these identities for the twining characters given in (3.14).

4. Conclusions

In this paper we have provided strong evidence for the conjecture of Eguchi, Ooguri and Tachikawa [11] that the elliptic genus of K3 carries a natural action of the Mathieu group $\mathbb{M}_{24}$. In particular, we have shown that the elliptic genus with the insertion of a group element in $\mathbb{M}_{24}$ leads, in the NS-sector, to a twining character that possesses nice modular properties. By considering the twining characters of symmetric power theories we have furthermore shown that the twining characters satisfy beautiful replication identities similar to those appearing for the McKay-Thompson series in Monstrous Moonshine. While the analogy of the twining characters with the McKay-Thompson series is quite striking, there are also marked differences: in the current context the twining characters are not Hauptmoduls for a genus zero congruence subgroup, and hence the replication relations contain also 'inhomogeneous' terms (for example the term proportional to $T_{[12\mathrm{D}]}$ in (3.22)).

So far we have only worked out about half of the twining characters in detail. In order to get a more complete understanding of the Mathieu action on the K3 elliptic genus, it would be very desirable to find closed form expressions for all twining characters. Probably some more refined techniques (rather than just guessing and confirming the answer) will be necessary; in particular, the techniques developed by Bantay & Gannon [1, 2] should be useful for this.

The above analysis suggests that the space of $\mathcal{N}=2$ half-BPS states of K3 carries an action of the Mathieu group $\mathbb{M}_{24}$. It is believed that these $\mathcal{N}=2$ half-BPS states form a subset of the quarter-BPS states (with respect to $\mathcal{N}=4$), for which the spectrum generating algebra is the so-called BPS-algebra of Harvey & Moore [19, 7]. The above group action on the spectrum then suggests that the Mathieu group $\mathbb{M}_{24}$ should be (contained in) the automorphism group of that Borcherds algebra, in complete analogy to what happened for Monstrous Moonshine. It would be very interesting to explore this idea in more detail.

Acknowledgments

We thank Terry Gannon for useful communications and Daniel Persson for comments on the draft. The research of MRG is partially supported by a grant from the Swiss National Science Foundation, and the research of RV is supported by an INFN Fellowship.

Note added: As this paper was being finalised, a preprint appeared `arXiv:1005.5415` [4] that contains overlapping results with those of section 3.1 of the present paper.

A. Definitions

Our conventions for the Dedekind eta and the Jacobi theta functions are

$$\begin{aligned}
\eta(\tau) &= q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n) \\
\vartheta_1(\tau,z) &= -iq^{\frac{1}{8}}y^{\frac{1}{2}} \prod_{n=1}^{\infty}(1-q^n)(1-yq^n)(1-y^{-1}q^{n-1}) \\
\vartheta_2(\tau,z) &= 2\,q^{\frac{1}{8}}\cos(\pi z) \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^n)(1+y^{-1}q^n) \\
\vartheta_3(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1+yq^{n-1/2})(1+y^{-1}q^{n-1/2}) \\
\vartheta_4(\tau,z) &= \prod_{n=1}^{\infty}(1-q^n)\,(1-yq^{n-1/2})(1-y^{-1}q^{n-1/2}) \quad .
\end{aligned} \tag{A.1}$$

Under modular transformations the ϑ and η functions transform as

$$\eta(\tau+1) = e^{\frac{2\pi i}{24}}\,\eta(\tau) \qquad \vartheta_2(\tau+1,z) = e^{\frac{2\pi i}{8}}\,\vartheta_2(\tau,z)\ , \tag{A.2}$$

$$\vartheta_3(\tau+1,z) = \vartheta_4(\tau,z) \qquad \vartheta_4(\tau+1,z) = \vartheta_3(\tau,z)\ , \tag{A.3}$$

as well as

$$\eta(-\tfrac{1}{\tau}) = (-i\tau)^{\frac{1}{2}}\,\eta(\tau) \qquad \vartheta_2(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_4(\tau,z)\ , \tag{A.4}$$

$$\vartheta_3(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_3(\tau,z) \qquad \vartheta_4(-\tfrac{1}{\tau},\tfrac{z}{\tau}) = (-i\tau)^{\frac{1}{2}}\,e^{\frac{i\pi z^2}{\tau}}\,\vartheta_2(\tau,z)\ . \tag{A.5}$$

The theta functions $\vartheta_a(\tau)$ are defined as $\vartheta_a(\tau) \equiv \vartheta_a(\tau, z=0)$.

A.1 McKay-Thompson series

For completeness we give here the first few terms of the McKay-Thompson series that appear in our analysis

$$T_{[8\mathrm{D}]} = q^{-1} + 4q + 2q^3 + 8q^5 - q^7 + \cdots \tag{A.6}$$

$$T_{[8\mathrm{E}]} = q^{-1} + 4q + 2q^3 - 8q^5 - q^7 + \cdots \tag{A.7}$$

$$T_{[8\mathrm{a}]} = q^{-1} - 20q - 62q^3 - 216q^5 - 641q^7 - 1636q^9 - \cdots \tag{A.8}$$

$$T_{[12\mathrm{D}]} = q^{-1} + 8q^2 + 28q^5 + 64q^8 + 134q^{11} + \cdots \tag{A.9}$$

$$T_{[12\mathrm{I}]} = q^{-1} + 2q + q^3 - 2q^7 - 2q^9 + 2q^{11} + \cdots \tag{A.10}$$

$$T_{[16\mathrm{A}]} = q^{-1} + 4q + 10q^3 + 24q^5 + 47q^7 + 84q^9 + \cdots \tag{A.11}$$

$$T_{[16\mathrm{B}]} = q^{-1} + 2q^3 - q^7 - 2q^{11} + 3q^{15} + 2q^{19} + \cdots \tag{A.12}$$

$$T_{[16\mathrm{b}]} = q^{-1} + 4q - 2q^3 + 8q^5 - q^7 + 20q^9 + \cdots \tag{A.13}$$

$$T_{[16\mathrm{d}]} = T_{[8\mathrm{D}]}(2\tau)^{1/2} = q^{-1} - 2q^3 - q^7 + 2q^{11} + \cdots \tag{A.14}$$

$$T_{[20\mathrm{C}]} = q^{-1} + q - 2q^2 + 2q^3 + 2q^4 - q^5 - 4q^7 + \cdots \tag{A.15}$$

$$T_{[24\mathrm{H}]} = q^{-1} + 2q + 5q^3 + 8q^5 + 14q^7 + 22q^9 + \cdots \tag{A.16}$$

$$T_{[24\mathrm{c}]} = q^{-1} - 2q + q^3 - 2q^7 + 2q^9 + 2q^{11} + \cdots \tag{A.17}$$

$$T_{[28\mathrm{B}]} = q^{-1} + 3q + 4q^2 + 9q^3 + 12q^4 + 15q^5 + 24q^6 + 39q^7 + \cdots \tag{A.18}$$

$$T_{[40\mathrm{a}]} = q^{-1} + 3q^3 + 4q^5 + 4q^7 + 4q^9 + 7q^{11} + \cdots \ , \tag{A.19}$$

see [15] and `http://www.research.att.com/ njas/sequences/` for more information about these series.

A.2 Twining characters for K3

Here we give the low order expansion of the remaining twining characters. Since the spectrum of the BPS states is real, the twining characters of conjugacy classes that only differ on conjugate representations agree. The remaining functions are therefore

$$\begin{aligned}\chi_{3\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 4q^{\frac{1}{2}} + 4q + 4q^2 + 8q^{\frac{5}{2}} + 2q^3 + 8q^{\frac{7}{2}} + 12q^4 + 4q^{\frac{9}{2}} + 16q^5 + 16q^{\frac{11}{2}} \\ &+ 5q^6 + 26q^{\frac{13}{2}} + 21q^7 + \cdots\Big)\end{aligned} \tag{A.20}$$

$$\begin{aligned}\chi_{6\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 4q^{\frac{1}{2}} + 8q + 16q^{\frac{3}{2}} + 32q^2 + 56q^{\frac{5}{2}} + 94q^3 + +152q^{\frac{7}{2}} + 240q^4 + 372q^{\frac{9}{2}} \\ &+ 560q^5 + 832q^{\frac{11}{2}} + 1197q^6 + 1594q^{\frac{13}{2}} + 1849q^7 + \cdots\Big)\end{aligned} \tag{A.21}$$

$$\begin{aligned}\chi_{7\mathrm{A}}(\tau) = \chi_{7\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + q^{\frac{1}{2}} + \tfrac{9}{2}q + 8q^{\frac{3}{2}} + \tfrac{27}{2}q^2 + 23q^{\frac{5}{2}} + \tfrac{81}{2}q^3 + 68q^{\frac{7}{2}} + 104q^4 \\ &+ 154q^{\frac{9}{2}} + \tfrac{469}{2}q^5 + 352q^{\frac{11}{2}} + \tfrac{987}{2}q^6 + \tfrac{1293}{2}q^{\frac{13}{2}} + 751q^7 + \cdots\Big)\end{aligned} \tag{A.22}$$

$$\begin{aligned}\chi_{8\mathrm{A}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 2q^{\frac{1}{2}} + 6q + 12q^{\frac{3}{2}} + 23q^2 + 42q^{\frac{5}{2}} + 74q^3 + 124q^{\frac{7}{2}} + 203q^4 + 324q^{\frac{9}{2}} \\ &+ 502q^5 + 768q^{\frac{11}{2}} + 1141q^6 + 1567q^{\frac{13}{2}} + 1866q^7 + \cdots\Big)\end{aligned} \tag{A.23}$$

$$\begin{aligned}\chi_{10\mathrm{A}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 4q^{\frac{1}{2}} + 7q + 12q^{\frac{3}{2}} + 24q^2 + 40q^{\frac{5}{2}} + 63q^3 + 100q^{\frac{7}{2}} + 153q^4 + 232q^{\frac{9}{2}} \\ &+ 342q^5 + 492q^{\frac{11}{2}} + 692q^6 + 900q^{\frac{13}{2}} + 1018q^7 + \cdots\Big)\end{aligned} \tag{A.24}$$

$$\begin{aligned}\chi_{11\mathrm{A}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 2q^{\frac{1}{2}} + 4q + 4q^{\frac{3}{2}} + 8q^2 + 14q^{\frac{5}{2}} + 17q^3 + 24q^{\frac{7}{2}} + 37q^4 + 52q^{\frac{9}{2}} \\ &+ 68q^5 + 88q^{\frac{11}{2}} + 116q^6 + 141q^{\frac{13}{2}} + 147q^7 + \cdots\Big)\end{aligned} \tag{A.25}$$

$$\begin{aligned}\chi_{12\mathrm{A}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 4q^{\frac{1}{2}} + 7q + 12q^{\frac{3}{2}} + 26q^2 + 48q^{\frac{5}{2}} + 78q^3 + 128q^{\frac{7}{2}} + 211q^4 + 336q^{\frac{9}{2}} \\ &+ 516q^5 + 780q^{\frac{11}{2}} + 1157q^6 + 1592q^{\frac{13}{2}} + 1886q^7 + \cdots\Big)\end{aligned} \tag{A.26}$$

$$\begin{aligned}\chi_{14\mathrm{A}}(\tau) = \chi_{14\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 3q^{\frac{1}{2}} + \tfrac{11}{2}q + 8q^{\frac{3}{2}} + \tfrac{33}{2}q^2 + 29q^{\frac{5}{2}} + \tfrac{87}{2}q^3 + 68q^{\frac{7}{2}} + 108q^4 \\ &+ 166q^{\frac{9}{2}} + \tfrac{487}{2}q^5 + 352q^{\frac{11}{2}} + \tfrac{1009}{2}q^6 + \tfrac{1335}{2}q^{\frac{13}{2}} + 761q^7 + \cdots\Big)\end{aligned} \tag{A.27}$$

$$\begin{aligned}\chi_{15\mathrm{A}}(\tau) = \chi_{15\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 3q^{\frac{1}{2}} + 6q + 10q^{\frac{3}{2}} + \tfrac{41}{2}q^2 + 37q^{\frac{5}{2}} + \tfrac{119}{2}q^3 + 96q^{\frac{7}{2}} + \tfrac{311}{2}q^4 \\ &+ 244q^{\frac{9}{2}} + 367q^5 + 544q^{\frac{11}{2}} + 795q^6 + \tfrac{2149}{2}q^{\frac{13}{2}} + 1250q^7 + \cdots\Big)\end{aligned} \tag{A.28}$$

$$\begin{aligned}\chi_{21\mathrm{A}}(\tau) = \chi_{21\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 4q^{\frac{1}{2}} + \tfrac{15}{2}q + 14q^{\frac{3}{2}} + \tfrac{57}{2}q^2 + 50q^{\frac{5}{2}} + \tfrac{165}{2}q^3 + 134q^{\frac{7}{2}} + 215q^4 \\ &+ 340q^{\frac{9}{2}} + \tfrac{1033}{2}q^5 + 772q^{\frac{11}{2}} + \tfrac{2271}{2}q^6 + 1545q^{\frac{13}{2}} + 1813q^7 + \cdots\Big)\end{aligned} \tag{A.29}$$

$$\begin{aligned}\chi_{23\mathrm{A}}(\tau) = \chi_{23\mathrm{B}}(\tau) = q^{-\frac{1}{4}}\Big(&1 + 3q^{\frac{1}{2}} + 7q + 14q^{\frac{3}{2}} + 26q^2 + 43q^{\frac{5}{2}} + \tfrac{149}{2}q^3 + 124q^{\frac{7}{2}} + \tfrac{383}{2}q^4 \\ &+ 296q^{\frac{9}{2}} + \tfrac{909}{2}q^5 + 682q^{\frac{11}{2}} + 985q^6 + \tfrac{2639}{2}q^{\frac{13}{2}} + 1548q^7 + \cdots\Big)\,.\end{aligned} \tag{A.30}$$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1A	2A	3A	5A	4B	7A	7B	8A	6A	11A	15A	15B	14A	14B	23A	23B	12B	6B	4C	3B	2B	10A	21A	21B	4A	12A
				2A			4B	3A		3A	3A	2A	2A			6B	3B	2B			2B	3B	3B	2A	6A
								2A								4C	2B								3A
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	7	5	3	3	2	2	1	1	1	0	0	0	0	0	0	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
252	28	9	2	4	0	0	0	1	−1	−1	−1	0	0	−1	−1	0	0	0	0	12	2	0	0	4	1
253	13	10	3	1	1	1	−1	−2	0	0	0	−1	−1	0	0	1	1	1	1	−11	−1	1	1	−3	0
1771	−21	16	1	−5	0	0	−1	0	0	1	1	0	0	0	0	−1	−1	−1	7	11	1	0	0	3	0
3520	64	10	0	0	−1	−1	0	−2	0	0	0	1	1	1	1	0	0	0	−8	0	0	−1	−1	0	0
45	−3	0	0	1	e_7^+	e_7^-	−1	0	1	0	0	$-e_7^+$	$-e_7^-$	−1	−1	1	−1	1	3	5	0	e_7^-	e_7^+	−3	0
45	−3	0	0	1	e_7^-	e_7^+	−1	0	1	0	0	$-e_7^-$	$-e_7^+$	−1	−1	1	−1	1	3	5	0	e_7^+	e_7^-	−3	0
990	−18	0	0	2	e_7^+	e_7^-	0	0	0	0	0	e_7^+	e_7^-	1	1	1	−1	−2	3	−10	0	e_7^-	e_7^+	6	0
990	−18	0	0	2	e_7^-	e_7^+	0	0	0	0	0	e_7^-	e_7^+	1	1	1	−1	−2	3	−10	0	e_7^+	e_7^-	6	0
1035	−21	0	0	3	$2e_7^+$	$2e_7^-$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^-$	$-e_7^+$	3	0
1035	−21	0	0	3	$2e_7^-$	$2e_7^+$	−1	0	1	0	0	0	0	0	0	−1	1	−1	−3	−5	0	$-e_7^+$	$-e_7^-$	3	0
1035	27	0	0	−1	−1	−1	1	0	1	0	0	−1	−1	0	0	0	2	3	6	35	0	−1	−1	3	0
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^+	e_{15}^-	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
231	7	−3	1	−1	0	0	−1	1	0	e_{15}^-	e_{15}^+	0	0	1	1	0	0	3	0	−9	1	0	0	−1	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^+	e_{23}^-	1	1	−2	−7	10	0	0	0	2	−1
770	−14	5	0	−2	0	0	0	1	0	0	0	0	0	e_{23}^-	e_{23}^+	1	1	−2	−7	10	0	0	0	2	−1
483	35	6	−2	3	0	0	−1	2	−1	1	1	0	0	0	0	0	0	3	0	3	−2	0	0	3	0
1265	49	5	0	1	−2	−2	1	1	0	0	0	0	0	0	0	0	0	−3	8	−15	0	1	1	−7	−1
2024	8	−1	−1	0	1	1	0	−1	0	−1	−1	1	1	0	0	0	0	0	8	24	−1	1	1	8	−1
2277	21	0	−3	1	2	2	−1	0	0	0	0	0	0	0	0	0	2	−3	6	−19	1	−1	−1	−3	0
3312	48	0	−3	0	1	1	0	0	1	0	0	−1	−1	0	0	0	−2	0	−6	16	1	1	1	0	0
5313	49	−15	3	−3	0	0	−1	1	0	0	0	0	0	0	0	0	0	−3	0	9	−1	0	0	1	1
5796	−28	−9	1	4	0	0	0	−1	−1	1	1	0	0	0	0	0	0	0	0	36	1	0	0	−4	−1
5544	−56	9	−1	0	0	0	0	1	0	−1	−1	0	0	1	1	0	0	0	0	24	−1	0	0	−8	1
10395	−21	0	0	−1	0	0	1	0	0	0	0	0	0	−1	−1	0	0	3	0	−45	0	0	0	3	0

Table 1: The character table of the Mathieu group. The rows correspond to the characters, while the columns describe the different conjugacy classes, numbered from 1 to 26 as in [11]. The second line gives the name of the conjugacy class that is used throughout the paper, where the number is the order of any element in the class. Where appropriate the classes of some powers are given in the third and fourth row, but only the non-trivial cases are stated. Finally, $e_p^\pm = (-1 \pm i\sqrt{p})/2$.

References

[1] P. Bantay and T. Gannon, *Conformal characters and the modular representation*, JHEP **0602**, 005 (2006) [`arXiv:hep-th/0512011`].

[2] P. Bantay and T. Gannon, *Vector-valued modular functions for the modular group and the hypergeometric equation*, Commun. Number Theory and Physics **1**, 637 (2008) [`arXiv:0705.2467 [math.NT]`]

[3] R. Borcherds, *Vertex algebras, Kac-Moody algebras and the Monster*, Proc. Nat. Acad. Sci. U.S.A. **83**, 3068 (1986).

[4] M.C.N. Cheng, *K3 surfaces, N=4 dyons, and the Mathieu group M24*, `arXiv:1005.5415 [hep-th]`.

[5] J.H. Conway and S. Norton, *Monstrous Moonshine*, Bull. Lond. Math. Soc. **11**, 308 (1979).

[6] R. Dijkgraaf, G.W. Moore, E.P. Verlinde and H.L. Verlinde, *Elliptic genera of symmetric products and second quantized strings*, Commun. Math. Phys. **185**, 197 (1997) [`arXiv:hep-th/9608096`].

[7] R. Dijkgraaf, E.P. Verlinde and H.L. Verlinde, *Counting dyons in N=4 string theory*, Nucl. Phys. B **484**, 543 (1997) [`arXiv:hep-th/9607026`].

[8] C.Y. Dong and G. Mason, *An orbifold theory of genus zero associated to the sporadic group M_{24}*, Commun. Math. Phys. **164**, 87 (1994).

[9] T. Eguchi and K. Hikami, *Superconformal algebras and Mock Theta Functions 2. Rademacher expansion for K3 surface*, Commun. Number Theory and Physics **3**, 531 (2009) [`arXiv:0904.0911 [math-ph]`].

[10] T. Eguchi, H. Ooguri, A. Taormina and S.K. Yang, *Superconformal algebras and string compactification on manifolds with $SU(N)$ holonomy*, Nucl. Phys. B **315**, 193 (1989).

[11] T. Eguchi, H. Ooguri and Y. Tachikawa, *Notes on the K3 surface and the Mathieu group M_{24}*, `arXiv:1004.0956 [hep-th]`.

[12] T. Eguchi and A. Taormina, *Character formulas for the $\mathcal{N} = 4$ superconformal algebra*, Phys. Lett. B **200**, 315 (1988).

[13] T. Eguchi and A. Taormina, *On the unitary representations of $N = 2$ and $N = 4$ superconformal algebras*, Phys. Lett. B **210**, 125 (1988).

[14] M. Eichler and D. Zagier, *The Theory of Jacobi Forms*, Birkhäuser (1985).

[15] D. Ford, J. McKay and S. Norton, *More on replicable functions*, Commun. Algebra **22**, 5175 (1994).

[16] I. Frenkel, J. Lepowski and A. Meurman, *Vertex Operator Algebras and the Monster*, Academic Press (1986).

[17] M.R. Gaberdiel, S. Gukov, C.A. Keller, G.W. Moore and H. Ooguri, *Extremal N=(2,2) 2d conformal field theories and constraints of modularity*, Commun. Number Theory and Physics **2**, 743 (2008) [`arXiv:0805.4216 [hep-th]`].

[18] T. Gannon, *Moonshine beyond the Monster: The Bridge connecting Algebra, Modular Forms and Physics*, Cambridge University Press (2006).

[19] J.A. Harvey and G.W. Moore, *Algebras, BPS states, and strings*, Nucl. Phys. B **463**, 315 (1996) [`arXiv:hep-th/9510182`]
J.A. Harvey and G.W. Moore, *On the algebras of BPS states*, Commun. Math. Phys. **197**, 489 (1998) [`arXiv:hep-th/9609017`].

[20] T. Kawai, Y. Yamada, and S.-K. Yang, *Elliptic genera and* $\mathcal{N} = 2$ *superconformal field theory*, Nucl. Phys. B **414**, 191 (1994) [`hep-th/9306096`].

[21] J.G. Thompson, *Some numerology between the Fischer-Griess Monster and the elliptic modular function*, Bull. Lond. Math. Soc. **11**, 352 (1979).

[22] M.P. Tuite, *Monstrous Moonshine from orbifolds*, Commun. Math. Phys. **146**, 277 (1992).

arXiv:1211.3703v3 [math.NT] 11 Dec 2013

The McKay-Thompson series of Mathieu Moonshine modulo two

Thomas Creutzig*, Gerald Höhn† and Tsuyoshi Miezaki‡§

Abstract

In this note, we describe the parity of the coefficients of the McKay-Thompson series of Mathieu moonshine. As an application, we prove a conjecture of Cheng, Duncan and Harvey stated in connection with umbral moonshine for the case of Mathieu moonshine.

1 Introduction

In 2010, Eguchi, Ooguri, and Tachikawa [9] discovered a phenomenon connecting the Mathieu group M_{24} and the elliptic genus of a K3 surface. To describe their observation, we let $q = e^{2\pi i\tau}$ and consider the function

$$\begin{aligned}\Sigma(\tau) &= -8(\,\mu(1/2;\tau) + \mu(\tau/2;\tau) + \mu((\tau+1)/2;\tau))\\ &= q^{-\frac{1}{8}}(-2 + 90q + 462q^2 + 1540q^3 + 4554q^4 + 11592q^5 + 27830q^6 + \cdots)\end{aligned}$$

where

$$\begin{aligned}\mu(z;\tau) &= \frac{ie^{\pi iz}}{\theta_1(z;\tau)}\sum_{n\in\mathbb{Z}}(-1)^n\,\frac{q^{\frac{1}{2}n(n+1)}e^{2\pi inz}}{1-q^ne^{2\pi iz}},\\ \theta_1(z;\tau) &= \sum_{n=-\infty}^{\infty} q^{\frac{1}{2}(n+\frac{1}{2})^2}e^{2\pi i(n+\frac{1}{2})(z+\frac{1}{2})}.\end{aligned}$$

---

[*] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada. email: creutzig@ualberta.ca

[†] Kansas State University, 138 Cardwell Hall, Manhattan, KS 66506-2602, USA. email: gerald@math.ksu.edu

[‡] Yamagata University, 1-4-12 Kojirakawa, Yamagata 990-8560, Japan. email: miezaki@e.yamagata-u.ac.jp

[§] This work was supported by JSPS KAKENHI Grant Number 22840003, 24740031.

The function $\Sigma(\tau)$ is a Mock modular form and counts the decomposition of the elliptic genus of a K3 surface into representations of the $\mathcal{N}$=4 Virasoro algebra.

The Mathieu moonshine phenomenon is that the first five coefficients appearing in the Fourier expansion of $\Sigma(\tau)$ divided by 2,

$$\{45, 231, 770, 2277, 5796\},$$

are equal to dimensions of irreducible representations of M_{24} and further coefficients can be written as simple sums of dimensions of the irreducible representations of M_{24}, for example $13915 = 3520 + 10395$. The reason for this mysterious phenomenon is still unknown.

This observation suggested the existence of a virtual graded M_{24}-module $K = \bigoplus_{n=-1}^{\infty} K_n \, q^{n/8}$ such that for $n \geq 0$ the K_n are honest M_{24}-representations. In analogy to the monstrous moonshine case [5], one can consider for an element g in the conjugacy class ℓX of M_{24} the so-called McKay-Thompson series

$$\Sigma_{\ell X}(\tau) = \sum_{n=-1}^{\infty} \mathrm{Tr}(g|K_n) \, q^{n/8}.$$

In [12, 7] (cf. also [1, 11, 2]), candidates for the 26 McKay-Thompson series for the Mathieu moonshine have been proposed. We note that the McKay-Thompson series for the M_{24} conjugacy classes in the pairs $(7A, 7B)$, $(14A, 14B)$, $(21A, 21B)$, $(15A, 15B)$ and $(23A, 23B)$ are equal to each other and we denote these cases shortly by ℓAB. We list in the appendix the 21 different McKay-Thompson series. Using explicit formulas, it can be shown that all the McKay-Thompson series have integer coefficients.

Elementary character theory [19] implies that the $\Sigma_{\ell X}(\tau)$ together uniquely determine the M_{24}-module K if it exists. Terry Gannon [13] showed that this is indeed the case.

Theorem 1.1 (Mathieu moonshine module)**.** *The McKay-Thompson series as in [12, 7] determine a virtual graded M_{24}-module $K = \bigoplus_{n=-1}^{\infty} K_n \, q^{n/8}$. For $n \geq 0$, the K_n are honest (and not only virtual) M_{24}-representations. Furthermore, K_n can be decomposed as a direct sum of M_{24}-representations of the form $\lambda \oplus \bar{\lambda}$ where λ is irreducible.*

This implies that if an irreducible representation λ is real, i.e. $\lambda \cong \bar{\lambda}$, then the multiplicity of an irreducible λ in K_n is even.

Unlike for the monstrous moonshine case where a vertex operator algebra structure was constructed in [10], there are yet no known underlying algebraic structures on the Mathieu moonshine module K.

To illustrate our main result, we consider the McKay-Thompson series for $7AB$:

$$\begin{aligned}\Sigma_{7AB}(\tau) =& \frac{1}{8}\,\left(\Sigma(\tau)\eta(\tau)^3 - 14\,\phi_2^{(7)}(\tau)\right)\Big/\,\eta(\tau)^3\\ =& -2q^{-1/8} - q^{7/8} + 4q^{31/8} - 2q^{47/8} + 2q^{55/8} - 3q^{63/8} + 6q^{87/8} - 6q^{103/8}\\ & -4q^{119/8} + 8q^{143/8} - 6q^{159/8} + 4q^{167/8} - 7q^{175/8} + 12q^{199/8} + \cdots .\end{aligned}$$

One observes that the coefficient of $q^{n/8}$ in $\Sigma_{7AB}(\tau)$ is odd if $n = 7m^2$, where m is odd. In general we show:

Theorem 1.2. *For a conjugacy class ℓX of M_{24}, the coefficient of $q^{n/8}$ in $\Sigma_{\ell X}(\tau)$ is odd if and only if $\ell X \in \{7AB, 14AB, 15AB, 23AB\}$ and $n = \ell m^2$, where m is odd, or $\ell X = 21AB$ and $n = \ell m^2$, where m is odd and not divisible by 3.*

For congruences of the Fourier coefficients of $\Sigma_{\ell X}(\tau)$ for other primes, we refer to the references [15, 17, 16]. One reason for considering the parity of the Fourier coefficients is that it explains the appearance of certain irreducible representations of M_{24}. The following conjecture was made in [3], which we state for the case of the Mathieu moonshine only.

Conjecture 1.1 ([3], Conj. 5.11)**.** *Let $n = \ell m^2 \equiv 7 \pmod 8$. Then the M_{24}-representation K_n determined by the coefficients of $q^{n/8}$ of the McKay-Thompson series contains the following conjugate pairs of irreducible representations:*

- *For $\ell = 7$, one of the pairs (χ_3, χ_4), (χ_{12}, χ_{13}) or (χ_{15}, χ_{16});*
- *for $\ell = 15$, the pair (χ_5, χ_6);*
- *for $\ell = 23$, the pair (χ_{10}, χ_{11}).*

Here, χ_i where $1 \leq i \leq 26$ denotes the i-th irreducible representation as listed in the ATLAS [4].

The paper is organized as follows. In Section 2, we study the cases $7AB$, $14AB$, $15AB$, $21AB$, $23AB$. In Section 3, we study the remaining cases. In the final section, we use Theorem 1.2 to prove Conjecture 1.1.

Acknowledgments. The first two authors like to thank the Hausdorff Research Institute of Mathematics for the hospitality received during their stay for the Mathematical Physics trimester program. We also like to thank M. Cheng, T. Gannon and S. Hohenegger for discussions. Finally, we are grateful to the referees for the suggested improvements of the exposition.

2 The non-even cases

In this section, we prove Theorem 1.2 for the cases $7AB$, $14AB$, $15AB$, $21AB$ and $23AB$.

The strategy is to obtain the parity properties of $\Sigma_{\ell X}(\tau)$ from the parity properties of a modular form for a group $\Gamma_0(N)$ which in turn can be proven by an application of Sturm's theorem [20]. The theorem allows to obtain a divisibility property of the Fourier coefficients of a modular form of weight k for $\Gamma_0(N)$ by veryfing this property only for the first n coefficients where n is an explicit given bound depending only on N and k,

$$n \geq \frac{k}{12}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)].$$

We let

$$\eta(\tau) = q^{1/24} \prod_{m=1}^{\infty} (1 - q^m)$$

be the Dedekind η-function and consider for $N \geq 2$ the Eisenstein series

$$\begin{aligned} \phi_2^{(N)}(\tau) &= \frac{24}{N-1}\, q\, \frac{d}{dq} \log\left(\frac{\eta(N\tau)}{\eta(\tau)}\right) \\ &= 1 + \frac{24}{N-1} \sum_{k=1}^{\infty} \sigma_1(k)(q^k - Nq^{Nk}) \end{aligned} \tag{1}$$

of weight 2 for $\Gamma_0(N)$.

Proof of Theorem 1.2 for the case $7AB$. Define the function $f_m(\tau)$ as follows:

$$f_m(\tau) = \frac{1}{4}\Big(\vartheta_3(\frac{m\tau}{8}) - \vartheta_4(\frac{m\tau}{8})\Big)$$

where $\vartheta_3(\tau) = 1 + \sum_{m=1}^{\infty} 2q^{m^2}$ and $\vartheta_4(\tau) = 1 + \sum_{m=1}^{\infty} 2(-q)^{m^2}$. Then

$$f_7(\tau) = \frac{1}{4}\Big(\vartheta_3(\frac{7\tau}{8}) - \vartheta_4(\frac{7\tau}{8})\Big) = q^{7/8} + q^{63/8} + q^{175/8} + q^{343/8} + \cdots$$

We call $f_7(\tau)$ the "parity function."

We have to show that all coefficients of $\Sigma_{7AB}(\tau) + f_7(\tau)$ are even since the coefficient of $q^{n/8}$ in $f_7(\tau)$ is odd if and only if $n = 7m^2$, where m is odd. This follows if all coefficients of $(\Sigma_{7AB}(\tau) + f_7(\tau))\eta(\tau)^3$ are even since $\eta(\tau)^{-3}$ has integral coefficients.

Let

$$E_2(\tau) = 1 - 24\sum_{m=1}^{\infty} \sigma_1(m)q^m.$$

By [6, p. 43, Example 4], we have

$$\Sigma(\tau) = -2\,\left(E_2(\tau) + 24\sum_{n=1}^{\infty}\frac{(-1)^n n q^{\frac{1}{2}n(n+1)}}{1-q^n}\right)\Big/\,\eta(\tau)^3\,. \tag{2}$$

Thus

$$\begin{aligned}(\Sigma_{7AB}(\tau) + f_7(\tau))\eta(\tau)^3 &= \frac{1}{8}\left(\Sigma(\tau)\eta(\tau)^3 - 14\,\phi_2^{(7)}(\tau)\right) + \eta(\tau)^3 f_7(\tau)\\ &= -\frac{1}{4}\Big(E_2(\tau) + 24\sum_{n=1}^{\infty}\frac{(-1)^n n q^{\frac{1}{2}n(n+1)}}{1-q^n} + 7\phi_2^{(7)}(\tau)\Big) + \eta(\tau)^3 f_7(\tau).\end{aligned}$$

First, we observe that the constant term is even. Since all coefficients, except for the constant term, of the function

$$-\frac{1}{4}\Big(E_2(\tau) + 24\sum_{n=1}^{\infty}\frac{(-1)^n n q^{\frac{1}{2}n(n+1)}}{1-q^n}\Big)$$

are even, it is enough to show that the coefficients of the following function, again except for the constant term,

$$-\frac{7}{4}\phi_2^{(7)}(\tau) + \eta(\tau)^3 f_7(\tau) = -\frac{7}{4} - 6q - 24q^2 - 28q^3 - 44q^4 - 42q^5 + \cdots\,.$$

are even.

Define the "correction function"

$$\frac{7}{4}\vartheta_3(\tau)^4 = \frac{7}{4} + 14q + 42q^2 + 56q^3 + 42q^4 + 84q^5 + \cdots .$$

The coefficients of $\frac{7}{4}\vartheta_3(\tau)^4$ without the constant term are even, hence it is enough to show that the coefficients of the function

$$-\frac{7}{4}\phi_2^{(7)}(\tau) + \eta(\tau)^3 f_7(\tau) + \frac{7}{4}\vartheta_3(\tau)^4 = 8q + 18q^2 + 28q^3 - 2q^4 + 42q^5 + \cdots$$

are even. We are going to prove this using Sturm's theorem. The theta functions $\vartheta_3(\tau)$ and $\vartheta_4(\tau)$ can be expressed as a quotient of η-functions, namely

$$\vartheta_3(\tau) = \frac{\eta(2\tau)^5}{\eta(\tau)^2\eta(4\tau)^2} \qquad \text{and} \qquad \vartheta_4(\tau) = \frac{\eta(\tau)^2}{\eta(2\tau)}.$$

It follows using [18, Theorem 1.64] that $\eta(8\tau)^3 f_7(8\tau)$ is a modular form of weight 2 for $\Gamma_0(448)$ and hence

$$\begin{aligned}\sum_{m=1}^{\infty} a_{7AB}(m)q^m &:= -\frac{7}{4}\phi_2^{(7)}(8\tau) + \eta(8\tau)^3 f_7(8\tau) + \frac{7}{4}\vartheta_3(8\tau)^4 \\ &= 8q^8 + 18q^{16} + 28q^{24} + \cdots\end{aligned}$$

is also a modular form of weight 2 for $\Gamma_0(448)$. Using Sturm's theorem [20] (see also [18, Theorem 2.58]) and the fact that $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(448)] = 768$, the computer verification that $a_{7AB}(m) \equiv 0 \pmod 2$ for $m \leq 129$ shows $a_{7AB}(m) \equiv 0 \pmod 2$ for all m.

This completes the proof of the case $7AB$. ☐

Since the other cases can be handled in complete analogy, we collect the relevant information in Table 1. For the definition of $\Sigma_{23AB}(\tau)$ in Appendix A, we use functions $f_{23,1}(\tau)$ and $f_{23,2}(\tau)$ which are modular forms for $\Gamma_0(23)$, explicitly given in [8, Appendix A.1].

3 The even cases

In this section, we prove that all coefficients of the McKay-Thompson series for the remaining conjugacy classes of M_{24} are divisible by two.

Table 1: Data for the proofs in Section 2

ℓX	parity function	correction function	Γ	$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$	Sturm bound
$7AB$	$f_7(\tau)$	$\frac{7}{4}\vartheta_3(\tau)^4$	$\Gamma_0(448)$	768	129
$14AB$	$f_7(\tau)$	$\frac{23}{12}\phi_2^{(2)}(\tau)$	$\Gamma_0(448)$	768	129
$15AB$	$f_{15}(\tau)$	$\frac{23}{12}\phi_2^{(2)}(\tau)$	$\Gamma_0(960)$	2304	385
$21AB$	$f_7(\tau)$-$f_{63}(\tau)$	$2\vartheta_3(\tau)^4$	$\Gamma_0(4032)$	9216	1537
$23AB$	$f_{23}(\tau)$	$\frac{23}{12}\phi_2^{(2)}(\tau)$	$\Gamma_0(1472)$	2304	385

We remark that the case $1A$ is clear since $\Sigma_{1A}(\tau) = \Sigma(\tau)$ and equation (2). It is trivial to see that for the cases $2B$, $3B$, $4A$, $4C$, $6B$, $10A$, $12A$, $12B$, the coefficients of $\Sigma_{\ell X}(\tau)$ are even because $\Sigma_{\ell X}(\tau)$ is -2 times an η-product (see the appendix).

3.1 The cases $2A$, $3A$, $4B$, $5A$, $6A$, $8A$

We give the detailed proof for the case of $2A$.

Proof. The McKay-Thompson series for $2A$ is

$$\begin{aligned}\Sigma_{2A}(\tau) &= \frac{1}{3}\left(\Sigma(\tau)\eta(\tau)^3 - 4\phi_2^{(2)}(\tau)\right)\Big/\,\eta(\tau)^3\\ &= -2q^{-1/8} - 6q^{7/8} + 14q^{15/8} - 28q^{23/8} + 42q^{31/8} + \cdots .\end{aligned}$$

The coefficients of $\Sigma(\tau)\eta(\tau)^3$, except for the constant term, are divisible by 24 (cf. (2)). Moreover, also the coefficients of $\phi_2^{(2)}(\tau)$, again except for the constant term, are divisible by 24. The constant term of the function $\Sigma(\tau)\eta(\tau)^3 - 4\phi_2^{(2)}(\tau)$ is 6. Therefore, the coefficients of $\Sigma_{2A}(\tau)$ are divisible by two. □

The proof for the other cases is analogous. The relevant information can be read off from Table 2 together with the appendix.

3.2 The case $11A$

Finally, it remains to show that the Fourier coefficients of the McKay-Thompson series for the case $11A$ are divisible by two. Because of the extra term

Table 2: Data for the proofs in Section 3.1

ℓX	N	divisor of $\phi_2^{(N)}(\tau)-1$
$2A$	2	24
$3A$	3	12
$5A$	5	6
$4B$	2, 4	24, 8
$8A$	4, 8	8, 24/7
$6A$	2, 3, 6	24, 12, 24/5

$\frac{264}{5}(\eta(\tau)\eta(11\tau))^2$ in $\Sigma_{11A}(\tau)$, it is not obvious that the coefficients except for the constant term in the numerator for $\Sigma_{11A}(\tau)$ as given in the appendix are divisible by 24. Therefore, we proceed similar to the cases in Section 3.1.

Proof. The McKay-Thompson series for $11A$ is

$$\Sigma_{11A}(\tau)=\frac{1}{12}\left(\Sigma(\tau)\eta(\tau)^3-22\,\phi_2^{(11)}(\tau)+\frac{264}{5}(\eta(\tau)\eta(11\tau))^2\right)\Big/\;\eta(\tau)^3\,.$$

The coefficients of $\Sigma(\tau)\eta(\tau)^3$, except for the constant term are divisible by 24. We need to prove that the coefficients of the function:

$$22\,\phi_2^{(11)}(\tau)-\frac{264}{5}(\eta(\tau)\eta(11\tau))^2=22+264q^2+264q^3+264q^4+\cdots$$

are also (except for the constant term) divisible by 24. The coefficients of $22\phi_2^{(2)}(\tau)$ (except for the constant term) are divisible by 24, hence it is enough to show that the coefficients of the function

$$\begin{aligned}\sum_{m=1}^{\infty}a_{11A}(m)q^m &:= 22\phi_2^{(11)}(\tau)-\frac{264}{5}(\eta(\tau)\eta(11\tau))^2-22\,\phi_2^{(2)}(\tau)\\ &= -528q-264q^2-1848q^3-264q^4-2904q^5-1584q^6-3696q^7+\cdots\end{aligned}$$

are divisible by 24. Note that this function is a modular form for $\Gamma_0(22)$ [14, p. 130, Proposition 19]. Using Sturm's theorem [18, Theorem 2.58] again and the fact that $[\mathrm{SL}_2(\mathbb{Z}):\Gamma_0(22)]=36$, the verification that $a_{11A}(m)\equiv 0 \pmod{24}$ for $m\leq 7$ shows $a_{11A}(m)\equiv 0 \pmod{24}$ for all m. Therefore, the coefficients of $\Sigma_{11A}(\tau)$ are divisible by two. ☐

Sections 2 and 3 together prove Theorem 1.2.

4 Multiplicities of Irreducibles

Recall that χ_i denotes the i-th irreducible M_{24}-representation as in [4]. We will use Theorem 1.2 to prove the following result about the multiplicities of the M_{24}-representation $\lambda \oplus \bar{\lambda}$ with irreducible λ inside the Mathieu moonshine module constituents K_n:

Theorem 4.1. *Let $n = \ell m^2 \equiv 7 \pmod 8$ and let K_n be the degree n-part of the Mathieu moonshine module K. Then the following numbers are odd (and therefore positive):*

- *For $\ell = 7$, the multiplicity of $\chi_3 \oplus \chi_4$ in K_n plus the multiplicity of $\chi_{12} \oplus \chi_{13}$ in K_n;*
- *for $\ell = 7$ and m divisible by 3, the multiplicity of $\chi_{15} \oplus \chi_{16}$ in K_n;*
- *for $\ell = 15$, the multiplicity of $\chi_5 \oplus \chi_6$ in K_n;*
- *for $\ell = 23$, the multiplicity of $\chi_{10} \oplus \chi_{11}$ in K_n.*

Proof. Let

$$K_n = \bigoplus_{i=1}^{26} m_{\chi_i} \chi_i \tag{3}$$

be the decomposition of the M_{24}-representation K_n into its irreducible constituents, i.e. m_{χ_i} is the multiplicity in which χ_i occurs in K_n. Taking on both sides of (3) the trace for an element $g \in M_{24}$ we have

$$\mathrm{Tr}(g|K_n) = \sum_{i=1}^{26} m_{\chi_i} \, \mathrm{Tr}(g|\chi_i) \tag{4}$$

and $\mathrm{Tr}(g|K_n)$ is the coefficient of $q^{n/8}$ in $\Sigma_{\ell X}(\tau)$ with ℓX the conjugacy class of g. We note that if K_n is non-zero then n is odd, i.e. we only need to consider the cases $n = \ell m^2$ with m odd.

First, we consider the cases $\ell = 15$ and $\ell = 23$ and let g be an element of type $15A$ or $23A$, respectively. If n is of the form ℓm^2 with m odd, the left-hand side of (4) is odd by Theorem 1.2. For the right-hand side an inspection of the character table of M_{24} [4] shows that if $\lambda \neq \bar{\lambda}$ one has that $\mathrm{Tr}(g|\lambda) = \mathrm{Tr}(g|\bar{\lambda})$ is integral unless $(\lambda, \bar{\lambda}) = (\chi_5, \chi_6)$ or (χ_{10}, χ_{11}), respectively, in which case $\mathrm{Tr}(g|\lambda + \bar{\lambda}) = -1$. Using Theorem 1.1, it follows that $m_\lambda = m_{\bar{\lambda}}$ has then to be odd.

For the cases $\ell = 7$, we let g be an element of type $7A$. Here, only the characters (χ_3, χ_4) or (χ_{12}, χ_{13}) can provide an odd contribution to the right-hand side of (4) and so in total an odd number of those pairs has to appear. For m divisible by 3, we take in addition an element g of type $21A$. Here (χ_3, χ_4), (χ_{12}, χ_{13}) and (χ_{15}, χ_{16}) provide an odd contribution. Since the total number of pairs of type (χ_3, χ_4) and (χ_{12}, χ_{13}) is odd and by Theorem 1.2 the left-hand side of (4) is even, it follows that the number of pairs (χ_{15}, χ_{16}) is odd, too. $\square$

It is clear that Theorem 4.1 implies Conjecture 1.1.

The argument in the proof can also be applied to the elements g of type ℓX as studied in Section 3. One directly recovers Theorem 1.2 for those cases. However, Theorem 1.1 as stated is a refinement which uses parts of our original arguments in its proof.

We remark that it is likely that the methods in proving Theorem 1.2 and Theorem 4.1 could also be applied to the other cases of umbral moonshine.

A McKay-Thompson series

We list all McKay-Thompson series using the ATLAS [4] names for the conjugacy classes of M_{24}. We refer to Section 2 for the functions used.

$$
\begin{aligned}
\mathbf{1A}: &\quad \Sigma_{1A}(\tau) = \Sigma(\tau) \\
\mathbf{2A}: &\quad \Sigma_{2A}(\tau) = \tfrac{1}{3}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 4\,\phi_2^{(2)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{2B}: &\quad \Sigma_{2B}(\tau) = -2\ \eta(\tau)^5\big/\,\eta(2\tau)^4 \\
\mathbf{3A}: &\quad \Sigma_{3A}(\tau) = \tfrac{1}{4}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 6\,\phi_2^{(3)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{3B}: &\quad \Sigma_{3B}(\tau) = -2\ \eta(\tau)^3\big/\,\eta(3\tau)^2 \\
\mathbf{4A}: &\quad \Sigma_{4A}(\tau) = -2\ \eta(2\tau)^8\big/\,\big(\eta(\tau)^3\eta(4\tau)^4\big) \\
\mathbf{4B}: &\quad \Sigma_{4B}(\tau) = \tfrac{1}{6}\ \Big(\Sigma(\tau)\eta(\tau)^3 + 2\phi_2^{(2)}(\tau) - 12\,\phi_2^{(4)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{4C}: &\quad \Sigma_{4C}(\tau) = -2\ \eta(\tau)\eta(2\tau)^2\big/\,\eta(4\tau)^2 \\
\mathbf{5A}: &\quad \Sigma_{5A}(\tau) = \tfrac{1}{6}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 10\,\phi_2^{(5)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{6A}: &\quad \Sigma_{6A}(\tau) = \tfrac{1}{12}\ \Big(\Sigma(\tau)\eta(\tau)^3 + 2\,\phi_2^{(2)}(\tau) + 6\,\phi_2^{(3)}(\tau) - 30\,\phi_2^{(6)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{6B}: &\quad \Sigma_{6B}(\tau) = -2\ \eta(2\tau)^2\eta(3\tau)^2\big/\,\big(\eta(\tau)\eta(6\tau)^2\big) \\
\mathbf{7AB}: &\quad \Sigma_{7AB}(\tau) = \tfrac{1}{8}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 14\,\phi_2^{(7)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{8A}: &\quad \Sigma_{8A}(\tau) = \tfrac{1}{12}\ \Big(\Sigma(\tau)\eta(\tau)^3 + 6\,\phi_2^{(4)}(\tau) - 28\,\phi_2^{(8)}(\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{10A}: &\quad \Sigma_{10A}(\tau) = -2\ \eta(2\tau)\eta(5\tau)\big/\,\eta(10\tau) \\
\mathbf{11A}: &\quad \Sigma_{11A}(\tau) = \tfrac{1}{12}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 22\,\phi_2^{(11)}(\tau) + \tfrac{264}{5}\,\big(\eta(\tau)\eta(11\tau)\big)^2\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{12A}: &\quad \Sigma_{12A}(\tau) = -2\ \eta(4\tau)^2\eta(6\tau)^3\big/\,\big(\eta(2\tau)\eta(3\tau)\eta(12\tau)^2\big) \\
\mathbf{12B}: &\quad \Sigma_{12B}(\tau) = -2\ \eta(\tau)\eta(4\tau)\eta(6\tau)\big/\,\big(\eta(2\tau)\eta(12\tau)\big) \\
\mathbf{14AB}: &\quad \Sigma_{14AB}(\tau) = \tfrac{1}{24}\Big(\Sigma(\tau)\eta(\tau)^3 + \tfrac{2}{3}\,\phi_2^{(2)}(\tau) + 14\phi_2^{(7)}(\tau) - \tfrac{182}{3}\,\phi_2^{(14)}(\tau) + \\
&\qquad\qquad\qquad + 112\,\eta(\tau)\eta(2\tau)\eta(7\tau)\eta(14\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{15AB}: &\quad \Sigma_{15AB}(\tau) = \tfrac{1}{24}\Big(\Sigma(\tau)\eta(\tau)^3 + \tfrac{3}{2}\,\phi_2^{(3)}(\tau) + 5\,\phi_2^{(5)}(\tau) - \tfrac{105}{2}\,\phi_2^{(15)}(\tau) + \\
&\qquad\qquad\qquad + 90\,\eta(\tau)\eta(3\tau)\eta(5\tau)\eta(15\tau)\Big)\Big/\,\eta(\tau)^3 \\
\mathbf{21AB}: &\quad \Sigma_{21AB}(\tau) = -\tfrac{1}{3}\,\big(\,7\,\eta(7\tau)^3\big/\,\big(\eta(3\tau)\eta(21\tau)\big)\ -\ \eta(\tau)^3\big/\,\eta(3\tau)^2\,\big) \\
\mathbf{23AB}: &\quad \Sigma_{23AB}(\tau) = \tfrac{1}{24}\ \Big(\Sigma(\tau)\eta(\tau)^3 - 46\,\phi_2^{(23)}(\tau) + \tfrac{276}{11}\,f_{23,1}(\tau) + \tfrac{1932}{11}\,f_{23,2}(\tau)\Big)\Big/\,\eta(\tau)^3
\end{aligned}
$$

References

[1] M. C. N. Cheng, $K3$ surfaces, N=4 dyons, and the Mathieu group M_{24}, *Commun. Number Theory Phys.* **4** (2010), no. 4, 623–657.

[2] M. C. N. Cheng and J. F. R. Duncan On Rademacher Sums, the Largest Mathieu Group, and the Holographic Modularity of Moonshine, Commun. Num. Theor. Phys. **6** (2012) 697.

[3] M. C. N. Cheng, J. F. R. Duncan and J. A. Harvey, Umbral Moonshine, arXiv:1204.2779.

[4] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, R. A. Wilson, Atlas of finite groups, Oxford University Press, 1985.

[5] J. H. Conway and S. P. Norton, Monstrous Moonshine, *Bull. Lond. Math. Soc.* **11** (1979), 308–339.

[6] A. Dabholkar, S. Murthy and D. Zagier, Quantum black holes, wall crossing, and mock modular forms, arXiv:1208.4074.

[7] T. Eguchi and K. Hikami, Note on twisted elliptic genus of $K3$ surface, *Phys. Lett. B* **694** (2011), 446–455.

[8] T. Eguchi and K. Hikami, Twisted elliptic genus for $K3$ and Borcherds product, *Letters in Math. Phys.*, **102** (2012), 203–222, arXiv:1112.5928.

[9] T. Eguchi, H. Ooguri and Y. Tachikawa, Notes on the $K3$ surface and the Mathieu group M_{24}, *Exp. Math.* **20** (2011), 91–96.

[10] I. Frenkel, J. Lepowsky and A. Meurman, *Vertex operator algebras and the Monster.* Pure and Applied Mathematics, 134, Academic Press, Inc., Boston, MA, 1988.

[11] M. R. Gaberdiel, S. Hohenegger and R. Volpato, Mathieu twining characters for $K3$, *J. High Energy Phys.* (2010), no. 9, **058**, 20 pp.

[12] M. R. Gaberdiel, S. Hohenegger and R. Volpato, Mathieu Moonshine in the elliptic genus of $K3$, *J. High Energy Phys.* (2010), no. 10, **062**, 24 pp.

[13] T. Gannon, Much ado about Mathieu, preprint (2012), arXiv:1211.5531.

[14] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, Berlin/New York, 1984.

[15] T. Miezaki, On the Mathieu mock theta function, *Proc. Japan Acad. Ser. A Math. Sci.* **88** (2012), 28–30.

[16] T. Miezaki, Congruences on the Fourier coefficients for mock theta functions, which relate finite groups, preprint.

[17] T. Miezaki, M. Waldherr, Congruence of the Fourier coefficients of the Mathieu mock theta function, submitted.

[18] K. Ono, *The web of modularity: arithmetic of the coefficients of modular forms and q-series*, CBMS Regional Conference Series in Mathematics, vol. 102, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2004.

[19] J.-P. Serre, *Représentations linéaires des groupes finis,* Springer, New York, 1977.

[20] J. Sturm, On the congruence of modular forms, *Number theory (New York, 1984–1985)*, 275–280, Lecture Notes in Math., 1240, Springer, Berlin, 1987.

ANNALES SCIENTIFIQUES DE L'É.N.S.

G. GONZALEZ-SPRINBERG
J.-L VERDIER

Construction géométrique de la correspondance de McKay

Annales scientifiques de l'É.N.S. 4e série, tome 16, nº 3 (1983), p. 409-449

<http://www.numdam.org/item?id=ASENS_1983_4_16_3_409_0>

© Gauthier-Villars (Éditions scientifiques et médicales Elsevier), 1983, tous droits réservés.

L'accès aux archives de la revue « Annales scientifiques de l'É.N.S. » (http://www.elsevier.com/locate/ansens) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Ann. scient. Éc. Norm. Sup.,
4^e série, t. 16, 1983, p. 409 à 449

CONSTRUCTION GÉOMÉTRIQUE DE LA CORRESPONDANCE DE McKAY

PAR G. GONZALEZ-SPRINBERG ET J. L. VERDIER

SOMMAIRE

0. *Introduction et notations* 409
1. *Résultats préliminaires :* $K(\tilde{\mathscr{S}})$ *et* $R(G)$ 412
2. *Construction géométrique de la correspondance* 415
3. *Rappel sur les groupes engendrés par des réflexions; les modules* M_ρ. 419
4. *Résolution* G-*équivariante de* A_V; *l'opérateur* Θ. 425
5. *La série de Poincaré; graduations et matrices* 426
6. *Désingularisations minimales; démonstration du théorème* 2.2 440
7. *Corollaire : Caractérisation de la désingularisation minimale* 448
Références 449

0. Introduction et notations

Les objets concernés par cet article sont essentiellement les solides platoniciens, i. e. les polyèdres réguliers de $\mathbb{R}^3$. La première démonstration du fait qu'ils sont au nombre de cinq (tétraèdre, octaèdre, cube, dodecaèdre et icosaèdre) est attribuée à Théétète ([P], [E]), qui a aussi prouvé qu'on peut les inscrire dans une sphère. Un groupe de rotations R qui laisse fixe un tel polyèdre est appelé groupe polyédral pur, et peut être considéré comme sous-groupe

© Gauthier-Villars

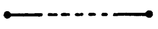
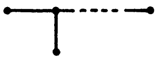
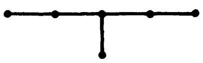
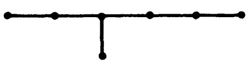
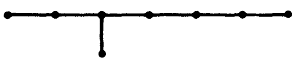
de PGL(2, $\mathbb{C}$), moyennant l'identification de la sphère S^2 avec la droite projective complexe $\mathbb{P}^1$. Le groupe G, image inverse de R par l'homomorphisme canonique de SL(2, $\mathbb{C}$) sur PGL(2, $\mathbb{C}$), est appelé groupe binaire polyédral. On trouve ainsi (à isomorphisme près) trois sous-groupes finis de SL(2, $\mathbb{C}$), car le cube et le dodecaèdre sont respectivement duaux de l'octaèdre et de l'icosaèdre par la dualité qui associe à un polyèdre celui dont les sommets sont les centres de ses faces (le tétraèdre est auto-dual), et deux polyèdres duaux ont le même groupe polyédral pur associé. Pour obtenir tous les sous-groupes finis de SL(2, $\mathbb{C}$) (à conjugaison près) il faut ajouter deux familles dénombrables : les groupes cycliques et les groupes diédraux binaires (*voir* par exemple [K] ou [D]); pour les groupes diédraux on peut aussi considérer les « groupes diédraux purs » comme étant les groupes de rotations de certains polyèdres appelés dièdres.

Soit G, donc, un sous-groupe fini de SL(2, $\mathbb{C}$), $G \neq \{1\}$; il agit naturellement sur $V = \mathbb{C}^2$; il a comme seul point fixe l'origine 0 et il agit librement sur $V - \{0\}$.

Posons $S = V/G$; notons $p : V \to S$ le morphisme de passage au quotient. La surface S a un point singulier isolé qui est un point double rationnel ([A], [Br], [D], [K]) et cette construction permet d'obtenir ainsi tous les points doubles rationnels [Br].

Soit $q : \tilde{S} \to S$ une désingularisation minimale de S, $D \subset \tilde{S}$ le diviseur exceptionnel; notons $\mathrm{Irr}(D) = \{d_1, \ldots, d_r\}$ l'ensemble des composantes irréductibles de D.

On peut associer à G un diagramme de Dynkin $\Gamma(G)$ en prenant le graphe dual de la résolution $\tilde{S}$, graphe dont les sommets sont en bijection avec Irr(D), et dont les arêtes joignant deux sommets représentent les points d'intersection des composantes de D correspondantes. Suivant le cas, le diagramme est donné par le tableau suivant :

G	$\lvert G \rvert$			$\Gamma(G)$
cyclique	$n+1$	$\mathbf{A}_n$:		
diédral binaire	$4(n-2)$	$\mathbf{D}_n$:		(n sommets, $n \geqq 1$)
tétraédral binaire	24	$\mathbf{E}_6$:		(n sommets, $n \geqq 4$)
octaédral binaire	48	$\mathbf{E}_7$:		
icosaédral binaire	120	$\mathbf{E}_8$:		

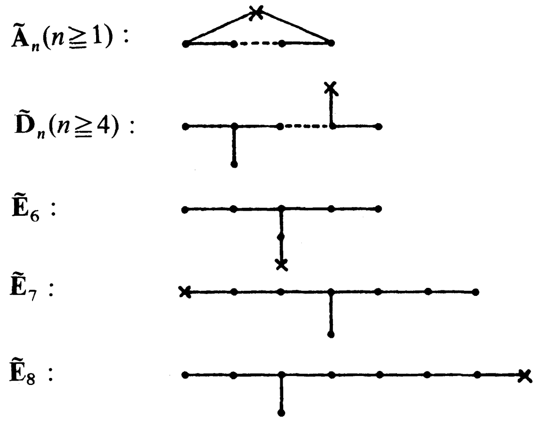
Pour tout sous-groupe fini G de SL(2, $\mathbb{C}$) on appelle $\Gamma(G)$ le type de G.

Considérons, par ailleurs, l'ensemble des classes d'équivalence de représentations non triviales de G : $\mathrm{Irr}(G)=\{\rho_1,\ldots,\rho_n\}$. Notons c la représentation canonique de G dans $\mathbb{C}^2$ donnée par l'inclusion $G \hookrightarrow SL(2, \mathbb{C})$, et ρ_0 la représentation triviale de degré 1. On définit la $(r+1)\times(r+1)$ matrice $C=((a_{ij}))$ par la formule de décomposition :

$$c\otimes\rho_i=\sum a_{ij}\,\rho_j,$$

où a_{ij} dénote la multiplicité de ρ_j dans $c\otimes\rho_i$.

J. McKay a remarqué que la matrice 2 Id – C (où Id est la matrice d'identité) est la matrice de Cartan du diagramme de Dynkin complété $\tilde{\Gamma}(G)$ de $\Gamma(G)$ ([M], Appendix III de [S1]). Du point de vue des représentations de groupes cette propriété a été analysée par J. Steinberg [St]. Ces diagrammes de Dykin complétés sont donc les suivants :



où on a indiqué par une croix le sommet ajouté.

La remarque de McKay permet de définir de façon purement combinatoire une bijection entre l'ensemble des classes d'équivalence de représentations irréductibles de G et l'ensemble de sommets de $\tilde{\Gamma}(G)$, qui induit une bijection entre le graphe de Irr(G) et celui de Irr(D) (en faisant correspondre à ρ_0 le sommet ajouté), bijection définie à automorphisme de diagramme près.

Cette bijection possède la propriété suivante : les degrés des représentations appartenant à Irr(G) coïncident avec les coefficients de la plus grande racine dans le système de racines de $\Gamma(G)$. Cette propriété peut être énoncée d'une autre façon, en considérant le cycle fondamental $\mathscr{Z}=\sum n_i d_i$ de $\tilde{S}$ (*voir* la définition dans [A]). Le cycle fondamental est (pour les singularités rationnelles) le cycle divisoriel défini par l'idéal (localement principal) $q^*(\mathscr{M}_S)$, avec $\mathscr{M}_S$ l'idéal maximal du point singulier. Soit $\mathscr{H}$ une section hyperplane générale de S contenant le point singulier. Alors l'image inverse $q^*(\mathscr{H})$ s'écrit $q^*(\mathscr{H})=\mathscr{Z}+\mathscr{T}$, où $\mathscr{T}$ est la transformée stricte et $\mathscr{Z}$ la partie de $q^*(\mathscr{H})$ à support dans D. On peut faire correspondre à $\mathscr{T}$ le sommet ajouté dans $\tilde{\Gamma}(G)$; on vérifie que les arêtes ajoutées correspondent aussi aux points d'intersection de $\mathscr{T}$ avec D. En ces termes, la bijection mentionnée associe à ρ_0 la transformée stricte $\mathscr{T}$ et à chaque $\rho\in\mathrm{Irr}(G)$ une composante $d\in\mathrm{Irr}(D)$ de telle sorte que le degré de ρ coïncide avec la multiplicité de d dans le cycle fondamental $\mathscr{Z}$.

Nous présentons ici une description géométrique de la correspondanxe de McKay. Plus précisément, à toute représentation irréductible non triviale ρ de G, on associe un fibré vectoriel $\mathscr{F}_\rho$ sur $\tilde{\mathscr{S}}$ (où $\tilde{\mathscr{S}}$ est la désingularisation minimale du germe $\mathscr{S}$ de la surface S au point singulier), fibré vectoriel dont la classe de Chern $c_1(\mathscr{F}_\rho)$ est un élément de la base canonique de $\operatorname{Pic}(\tilde{\mathscr{S}})$ en bijection avec les sommets de Irr(D) (¹).

1. Résultats préliminaires : $K(\tilde{\mathscr{S}})$ et R(G)

1.1. Soient $\mathscr{S}$ le germe de S au point singulier, $q : \tilde{\mathscr{S}} \to \mathscr{S}$ une résolution minimale, $\operatorname{Pic}(\tilde{\mathscr{S}})$ le groupe de Picard de $\tilde{\mathscr{S}}$ des classes d'isomorphisme de fibrés de rang 1 ou de $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules inversibles, et $\operatorname{Irr}(D) = \{ d_1, \ldots, d_r \}$ l'ensemble des composantes irréductibles du diviseur exceptionnel $D \subset \tilde{\mathscr{S}}$. Pour chaque $d \in \operatorname{Irr}(D)$, notons e_d la classe de $\mathcal{O}_{\tilde{\mathscr{S}}}(d)$ dans $\operatorname{Pic}(\tilde{\mathscr{S}})$, et $\operatorname{Pic}_D(\tilde{\mathscr{S}})$ le sous-groupe de $\operatorname{Pic}(\tilde{\mathscr{S}})$ engendré par les e_d. Soit $\operatorname{Pic}_D(\tilde{\mathscr{S}})^*$ le dual de $\operatorname{Pic}_D(\tilde{\mathscr{S}})$, et notons $\delta : \operatorname{Pic}(\tilde{\mathscr{S}}) \to \operatorname{Pic}_D(\tilde{\mathscr{S}})^*$ l'homomorphisme défini par la forme intersection de $\operatorname{Pic}(\tilde{\mathscr{S}})$. On sait que δ est un isomorphisme, car $\mathscr{S}$ est une singularité rationnelle [A]. Notons $\{ b_d \mid d \in \operatorname{Irr}(D) \}$ la base de $\operatorname{Pic}(\tilde{\mathscr{S}})$ transportée par δ^{-1} de la base duale de $\{ e_d \mid d \in \operatorname{Irr}(D) \}$. La matrice de l'injection canonique $\operatorname{Pic}_D(\tilde{\mathscr{S}}) \hookrightarrow \operatorname{Pic}(\tilde{\mathscr{S}})$ dans les bases $\{ e_d \}$ et $\{ b_d \}$ respectivement, est appelée la matrice d'intersection.

On considère d'autre part l'anneau de Grothendieck $K(\tilde{\mathscr{S}})$ des classes de $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules cohérents, et on note $K_D(\tilde{\mathscr{S}})$ le sous-groupe engendré par les classes de $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules à support dans D; $K_D(\tilde{\mathscr{S}})$ est un idéal de $K(\tilde{\mathscr{S}})$. Pour tout $\mathcal{O}_{\tilde{\mathscr{S}}}$-module cohérent M notons [M] sa classe dans $K(\tilde{\mathscr{S}})$, et pour tout $d \in \operatorname{Irr}(D)$ posons $[d] = [\mathcal{O}_d]$. Finalement notons $rg : K(\tilde{\mathscr{S}}) \to \mathbb{Z}$ [resp. $c_1 : K(\tilde{\mathscr{S}}) \to \operatorname{Pic}(\tilde{\mathscr{S}})$] l'application induite par celle qui associe à un $\mathcal{O}_{\tilde{\mathscr{S}}}$-module cohérent son rang (resp. sa première classe de Chern).

Avec les notations précédentes on a :

1.2. PROPOSITION. — (i) *L'application* $\varphi : K(\tilde{\mathscr{S}}) \to \mathbb{Z} \oplus \operatorname{Pic}(\tilde{\mathscr{S}})$ *définie par* $x \mapsto (rg(x), c_1(x))$ *est un isomorphisme de groupes commutatifs. Pour tout* $L \in \operatorname{Pic}(\tilde{\mathscr{S}})$ *on a* $\varphi([L]) = (1, L)$.

(ii) *Le noyau de* $rg : K(\tilde{\mathscr{S}}) \to \mathbb{Z}$ *est un idéal de carré nul.*

(iii) *L'image par* φ *de* $K_D(\tilde{\mathscr{S}})$ *est* $\operatorname{Pic}_D(\tilde{\mathscr{S}})$. *Pour tout* $d \in \operatorname{Irr}(D)$ *on a* $\varphi([d]) = (0, e_d)$.

Démonstration. — Si $L \in \operatorname{Pic}(\tilde{\mathscr{S}})$, l'égalité $\varphi([L]) = (1, L)$ résulte de la définition de φ; par suite, si $n \in \mathbb{Z}$ on a :

$$\varphi((n-1)\mathbf{1} + [L]) = (n, L),$$

d'où la surjectivité de φ.

L'injectivité résulte du fait que le noyau est le sous-groupe $K_2(\tilde{\mathscr{S}}) \subset K(\tilde{\mathscr{S}})$ des classes de modules à support de codimension ≧ 2 dont on vérifie immédiatement qu'il est nul; ceci

---

(¹) Ces résultats ont été annoncés en [G-V]. H. Knörrer nous a fait parvenir une lettre en décembre 1981, où il propose une démonstration différente, trouvée indépendamment.

montre aussi (ii). Finalement, soit $d \in \mathrm{Irr}(d)$. On a la suite exacte :

$$0 \to \mathcal{O}_{\tilde{\mathcal{S}}}(-d) \to \mathcal{O}_{\tilde{\mathcal{S}}} \to \mathcal{O}_d \to 0,$$

d'où (iii).

1.3. Soit R(G) l'anneau des représentations de G. Si ρ est une représentation de G, on note $\rho^{\vee}$ la représentation duale ou contragrédiente de ρ.

Notons $\mathcal{V}$ le germe de V à l'origine 0.

On considère l'anneau de Grothendieck $\mathrm{K}^{\mathrm{G}}(\mathcal{V})$ des classes de $\mathcal{O}_{\mathcal{V}}$-modules cohérents G-équivariants, et le sous-groupe $\mathrm{K}^{\mathrm{G}}(0)$ de $\mathrm{K}^{\mathrm{G}}(\mathcal{V})$ de classes de $\mathcal{O}_{\mathcal{V}}$-modules à support dans 0.

1.4. Proposition. — (i) *L'application qui à une représentation* ρ *de* G *dans un* ℂ*-espace vectoriel* E_ρ *associe le* $\mathcal{O}_{\mathcal{V}}$*-module* G*-équivariant* $\mathcal{O}_{\mathcal{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^{\vee}}$ *induit un isomorphisme d'anneaux* $\mathrm{R}(\mathrm{G}) \xrightarrow{\sim} \mathrm{K}^{\mathrm{G}}(\mathcal{V})$ *qui permet d'identifier ces anneaux.*

(ii) *Soit* $\mathcal{M}_{\mathcal{V}}$ *l'idéal maximal en l'origine* 0 *de* $\mathcal{V}$. *La classe* $[\mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}}]$ *dans* $\mathrm{K}^{\mathrm{G}}(\mathcal{V}) = \mathrm{R}(\mathrm{G})$ *est* $2-c$. *L'image de* $\mathrm{K}^{\mathrm{G}}(0)$ *dans* $\mathrm{K}^{\mathrm{G}}(\mathcal{V}) = \mathrm{R}(\mathrm{G})$ *est l'idéal engendré par* $2-c$.

(iii) *L'idéal* $(2-c)\,\mathrm{R}(\mathrm{G})$ *est un* ℤ*-module libre engendré par* $(2-c)\,\rho$, $\rho \in \mathrm{Irr}(\mathrm{G})$. *On a la relation :*

$$2-c = \sum_{\rho \in \mathrm{Irr}(\mathrm{G})} (\deg \rho).(c-2)\,\rho.$$

Démonstration. — L'affirmation de (i) résulte du fait que l'homomorphisme inverse est celui induit par l'application qui à un $\mathcal{O}_{\mathcal{V}}$-module G-équivariant M associe l'élément de R(G) défini par $\sum_i (-1)^i [\mathrm{Tor}_i^{\mathcal{O}_{\mathcal{V}}}(\mathrm{M}, \mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}})]$.

Pour démontrer l'égalité $[\mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}}] = 2-c$, on considère la suite exacte de Koszul de $\mathcal{O}_{\mathcal{V}}$-modules G-équivariants :

$$0 \to \Lambda^2 \mathrm{T}_{\mathcal{V}} \to \mathrm{T}_{\mathcal{V}} \to \mathcal{O}_{\mathcal{V}} \to \mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}} \to 0,$$

où $\mathrm{T}_{\mathcal{V}}$ est le faisceau tangent de $\mathcal{V}$.

Or $\mathrm{T}_{\mathcal{V}} \cong \mathcal{O}_{\mathcal{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_c$ car c est par définition la représentation canonique de G, représentation qui est auto-dual (i.e. $c = c^{\vee}$); et d'autre part $\Lambda^2 \mathrm{T}_{\mathcal{V}} \cong \mathcal{O}_{\mathcal{V}}$ car on a $\mathrm{G} \subset \mathrm{SL}(2, \mathbb{C})$. Par conséquent la suite exacte précédente fournit une résolution G-équivariante de $\mathcal{O}_{\mathcal{V}}$-modules qu'on peut écrire :

$$0 \to \mathcal{O}_{\mathcal{V}} \to \mathcal{O}_{\mathcal{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{c^{\vee}} \to \mathcal{O}_{\mathcal{V}} \to \mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}} \to 0,$$

d'où l'égalité qu'on voulait démontrer.

En considérant l'isomorphisme de R(G) sur $\mathrm{K}^{\mathrm{G}}(0)$ [analogue à celui de R(G) sur $\mathrm{K}^{\mathrm{G}}(\mathcal{V})$] induit par l'application qui à ρ associe $(\mathcal{O}_{\mathcal{V}}/\mathcal{M}_{\mathcal{V}}) \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^{\vee}}$, on a le diagramme suivant dont la commutativité résulte de la résolution précédente de $\mathcal{O}_{\mathscr{V}}/\mathcal{M}_{\mathscr{V}}$:

$$\begin{tikzcd}
\mathrm{R}(\mathrm{G}) \arrow[r, "2-c"] \arrow[d, "\cong"] & \mathrm{R}(\mathrm{G}) \arrow[d, "\cong"] \\
\mathrm{K}^{\mathrm{G}}(0) \arrow[r] & \mathrm{K}^{\mathrm{G}}(\mathscr{V})
\end{tikzcd}$$

Par suite l'image de $\mathrm{K}^{\mathrm{G}}(0)$ dans $\mathrm{K}^{\mathrm{G}}(\mathscr{V})=\mathrm{R}(\mathrm{G})$ est l'idéal engendré par $(2-c)$.

Pour démontrer (iii), on utilise le fait que la matrice $\mathrm{C}-2\,\mathrm{Id}$ [où $\mathrm{C}=((a_{ij}))$ est définie par $c\otimes\rho_i=\sum a_{ij}\,\rho_j$] restreinte au sous-groupe de $\mathrm{R}(\mathrm{G})$ engendré par les éléments de $\mathrm{Irr}(\mathrm{G})$ est définie négative, donc en particulier est non singulière. Par conséquent $(2-c)\,\mathrm{R}(\mathrm{G})$ est un $\mathbb{Z}$-module libre de rang au moins égal à $r=\mathrm{card}(\mathrm{Irr}(\mathrm{G}))$ et au plus égal à $r+1$, rang de $\mathrm{R}(\mathrm{G})$.

D'autre part, ainsi qu'on le voit en considérant les caractères correspondants, on a la relation $c\,\mathrm{reg}(\mathrm{G})=2\,\mathrm{reg}(\mathrm{G})$, où $\mathrm{reg}(\mathrm{G})=\sum_{i=0}^{r}(\deg\rho_i)\,\rho_i$ est la représentation régulière de G, car c est de degré 2. Comme $\mathrm{reg}(\mathrm{G})$ est un élément primitif de $\mathrm{R}(\mathrm{G})$, on a la suite exacte :

$$0\rightarrow\mathbb{Z}.\mathrm{reg}(\mathrm{G})\rightarrow\mathrm{R}(\mathrm{G})\xrightarrow{(2-c)}\mathrm{R}(\mathrm{G}),$$

et par suite $(2-c)\,\mathrm{R}(\mathrm{G})$ est de rang égal à r, et coïncide avec le $\mathbb{Z}$-module engendré par $(2-c)\,\rho$, $p\in\mathrm{Irr}(\mathrm{G})$. De la relation précédente on obtient :

$$2-c=\sum_{\rho\in\mathrm{Irr}(\mathrm{G})}(\deg\rho)(c-2)\,\rho,$$

car ρ_0 est la représentation triviale de degré 1; ce qui achève la démonstration de la proposition 1.4.

1.5. REMARQUE. — Le diagramme commutatif considéré dans la démonstration précédente se prolonge en le diagramme commutatif suivant :

$$\begin{tikzcd}
\mathrm{R}(\mathrm{G}) \arrow[r] \arrow[d, "\cong"] & \mathrm{R}(\mathrm{G}) \arrow[r] \arrow[d, "\cong"] & \mathrm{R}(\mathrm{G})/(2-c)\,\mathrm{R}(\mathrm{G}) \arrow[r] \arrow[d, "\cong"] & 0 \\
\mathrm{K}^{\mathrm{G}}(0) \arrow[r] & \mathrm{K}^{\mathrm{G}}(\mathscr{V}) \arrow[r] & \tilde{\mathrm{K}}^{\mathrm{G}}(\mathscr{V}-\{0\}) \arrow[r] & 0
\end{tikzcd}$$

où les lignes sont exactes et où $\tilde{\mathrm{K}}^{\mathrm{G}}(\mathscr{V}-\{0\})$ est le sous-groupe de Grothendieck construit à l'aide des faisceaux cohérents qui se prolongent à $\mathscr{V}$.

Ce diagramme permet de calculer $\tilde{\mathrm{K}}^{\mathrm{G}}(\mathscr{V}-\{0\})$: il est somme directe de son sous-groupe de torsion et de son sous-groupe cyclique infini engendré par son élément unité.

On obtient les résultat suivants :

Type de G :		$\tilde{K}^G(\mathscr{V}-\{0\})\cong$:
$\mathbf{A}_n (n \geqq 1)$		$\mathbb{Z}.1 \oplus \mathbb{Z}/(n+1)\mathbb{Z}$
$\mathbf{D}_n (n \geqq 4)$	si n pair	$\mathbb{Z}.1 \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$
	si n impair	$\mathbb{Z}.1 \oplus \mathbb{Z}/4\mathbb{Z}$
$\mathbf{E}_6$		$\mathbb{Z}.1 \oplus \mathbb{Z}/3\mathbb{Z}$
$\mathbf{E}_7$		$\mathbb{Z}.1 \oplus \mathbb{Z}/6\mathbb{Z}$
$\mathbf{E}_8$		$\mathbb{Z}.1$

2. Construction géométrique de la correspondance

2.1. On utilise les notations des paragraphes précédents.

Soit $\tilde{\mathscr{V}}=(\mathscr{V}\times_{\mathscr{S}}\tilde{\mathscr{S}})_{\text{réd}}$ le produit fibré réduit de $\mathscr{V}$ et $\tilde{\mathscr{S}}$ au-dessus de $\mathscr{S}$. On a le diagramme commutatif suivant :

$$\begin{tikzcd}
\tilde{\mathscr{V}} \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\mathrm{pr}_1"'] & \tilde{\mathscr{S}} \arrow[d, "q"] \\
\mathscr{V} \arrow[r, "p"'] & \mathscr{S}
\end{tikzcd}$$

où pr_1 et pr_2 dénotent les projections canoniques.

2.2. THÉORÈME. — *Soit* π : $\mathrm{R(G)}=\mathrm{K^G}(\mathscr{V})\to \mathrm{K}(\tilde{\mathscr{S}})$ *l'application induite par la composition* $\mathrm{Inv}\circ \mathrm{pr}_{2*}\circ \mathrm{pr}_1^*$, *où* pr_1^* *et* pr_{2*} *sont les morphismes canoniques et où* Inv *est l'application qui associe à un* $\mathcal{O}_{\tilde{\mathscr{S}}}$*-module* G-*équivariant* M *le sous-module* $\mathrm{M^G}$ *des invariants pour l'action de* G. *Alors on a :*

(i) *Pour tout* $\rho\in \mathrm{Irr(G)}$ *il existe un unique* $d_\rho \in \mathrm{Irr(D)}$ *tel que* $\pi(\rho)=(\deg \rho, b_{d_\rho})$. *L'application* $\rho \to d_\rho$ *de* Irr(G) *dans* Irr(D) *est bijective. Pour* $\rho_i \neq \rho_j$ *on a* $(d_{\rho_i}.d_{\rho_j})=a_{ij}$ *(où* a_{ij} *est la multiplicité de* ρ_j *dans* $c\otimes\rho_i$*).*

(ii) $\pi(2-c)$ *est la classe* $[\mathscr{Z}]$ *du cycle fondamental, et pour tout* $\rho\in \mathrm{Irr(G)}$ *on a* $\pi((c-2)\rho)=[d_\rho]$.

(iii) *L'application* π *est un isomorphisme de* $\mathbb{Z}$*-modules.*

(iv) $\pi((2-c)\,\mathrm{R(G)})=\mathrm{Pic_D}(\tilde{\mathscr{S}})$.

(v) *Par passage au quotient* π *induit un isomorphisme d'anneaux :*

$$\mathrm{R(G)}/(2-c)\,\mathrm{R(G)} \xrightarrow{\sim} \mathrm{K}(\tilde{\mathscr{S}})/\mathrm{K_D}(\tilde{\mathscr{S}}).$$

2.3. COROLLAIRE. — *On a :*

$$[\mathscr{Z}]=\sum_{\rho\in \mathrm{Irr(G)}} (\deg \rho)\,[d_\rho].$$

Démonstration du corollaire 2.3. — En appliquant π aux deux membres de l'égalité :

$$2-c=\sum_{\rho\in \mathrm{Irr}(G)} (\deg\rho)(c-2)\,\rho$$

[prop. 1.2 (iii)], on obtient l'égalité cherchée comme conséquence du théorème 2.2 (ii).

La proposition suivante est un résultat préliminaire pour la démonstration du théorème 2.2; ses hypothèses sont un peu moins restrictives que celles satisfaites par la surface $\mathscr{S}$ de ce théorème.

2.4. PROPOSITION. — *Soit* $p:(\mathscr{V},0)\to(\mathscr{S},\bar{0})$ *un morphisme fini dominant où* $(\mathscr{V},0)$ *est un germe de surface normale, et où* $(\mathscr{S},\bar{0})$ *est un germe de singularité rationnelle de surface.*

Soit $q:\tilde{\mathscr{S}}\to\mathscr{S}$ *une désingularisation de* $\mathscr{S}$; *notons* $\mathscr{V}'=\mathscr{V}\times_{\mathscr{S}}\tilde{\mathscr{S}}$ *le produit fibré de* $\mathscr{V}$ *et* $\tilde{\mathscr{S}}$ *au-dessus de* $\mathscr{S}$ *et* $\tilde{\mathscr{V}}=\mathscr{V}'_{\mathrm{red}}$ *le réduit de* $\mathscr{V}'$. *Alors on a :*

(i) *L'homomorphisme canonique* (*de* $\mathcal{O}_{\mathscr{S}}$-*complexes*) $\mathcal{O}_{\mathscr{V}}\to q_*\mathcal{O}_{\tilde{\mathscr{V}}}$ *est un isomorphisme et on a de plus* $\mathrm{R}^1 q_*\mathcal{O}_{\tilde{\mathscr{V}}}=0$.

(ii) *Le* $\mathcal{O}_{\tilde{\mathscr{S}}}$-*module* $\mathcal{O}_{\tilde{\mathscr{V}}}$ *est localement libre.*

Montrons d'abord un lemme :

2.5. LEMME. — *Soit* A *un anneau commutatif noetherien et intègre, et* B *une* A-*algèbre finie irréductible, dominante, et génériquement réduite. Notons* $\mathscr{N}$ *l'idéal des éléments nilpotents de l'anneau* B *et* $\mathrm{Tors}_{\mathrm{A}}\,\mathrm{B}$ *la* A-*torsion de* B. *Alors on a :*

$$\mathscr{N}=\mathrm{Tors}_{\mathrm{A}}\,\mathrm{B}.$$

Démonstration du lemme 2.5. — En effet $\mathscr{N}$ est un A-module de type fini génériquement nul donc $\mathscr{N}\subset\mathrm{Tors}_{\mathrm{A}}\,\mathrm{B}$. L'algèbre $\mathrm{B}_{\mathrm{red}}$ est intègre donc sans A-torsion, car dominante, d'où $\mathrm{Tors}_{\mathrm{A}}\,\mathrm{B}\subset\mathscr{N}$.

Démonstration de la proposition 2.4. — Donc on a par la formule de projection pour les morphismes non nécessairement plats [Gr] :

$$\mathcal{O}_{\mathscr{V}}\mathop{\overset{\mathbb{L}}{\otimes}}_{\mathcal{O}_{\mathscr{S}}}\mathrm{R}q_*(\mathcal{O}_{\tilde{\mathscr{S}}})=\mathrm{R}q_*\left(\mathcal{O}_{\tilde{\mathscr{S}}}\mathop{\overset{\mathbb{L}}{\otimes}}_{\mathcal{O}_{\mathscr{S}}}\mathcal{O}_{\mathscr{V}}\right)$$

où le $\overset{\mathbb{L}}{\otimes}$ dénote le produit tensoriel dans la catégorie dérivée.

Étant donné que $\mathscr{S}$ est une singularité rationnelle, on a :

$$\mathrm{R}q_*(\mathcal{O}_{\tilde{\mathscr{S}}})=\mathcal{O}_{\mathscr{S}},$$

et par suite :

$$\mathcal{O}_{\mathscr{V}}=\mathrm{R}q_*\left(\mathcal{O}_{\tilde{\mathscr{S}}}\mathop{\overset{\mathbb{L}}{\otimes}}_{\mathcal{O}_{\mathscr{S}}}\mathcal{O}_{\mathscr{V}}\right).$$

Donc on a une suite spectrale, dont le terme $\mathrm{E}_2^{r,s}$ est $\mathrm{R}^r q_*\left(\mathscr{H}^s\left(\mathcal{O}_{\tilde{\mathscr{S}}}\mathop{\overset{\mathbb{L}}{\otimes}}_{\mathcal{O}_{\mathscr{S}}}\mathcal{O}_{\mathscr{V}}\right)\right)$, qui aboutit à $\mathcal{O}_{\mathscr{V}}$.

Comme $R^r q_* = 0$ pour $r \neq 0, 1$, on a la suite exacte de $\mathcal{O}_{\mathscr{S}}$-module :

$$0 \to R^1 q_*(\mathrm{Tor}_1^{\mathcal{O}_{\mathscr{S}}}(\mathcal{O}_{\tilde{\mathscr{S}}}, \mathcal{O}_{\mathscr{V}})) \to \mathcal{O}_{\mathscr{V}} \to q_*(\mathcal{O}_{\mathscr{V}'}) \to 0,$$

et :

$$R^1 q_* \mathcal{O}_{\mathscr{V}'} = 0.$$

Or le support de $R^1 q_*(\mathrm{Tor}_1^{\mathcal{O}_{\mathscr{S}}}(\mathcal{O}_{\tilde{\mathscr{S}}}, \mathcal{O}_{\mathscr{V}}))$ est de codimension positive, donc $R^1 q_*(\mathrm{Tor}_1^{\mathcal{O}_{\mathscr{S}}}(\mathcal{O}_{\tilde{\mathscr{S}}}, \mathcal{O}_{\mathscr{V}})) = 0$, car $\mathcal{O}_{\mathscr{V}}$ est intègre, et par suite $\mathcal{O}_{\mathscr{V}} \to q_*(\mathcal{O}_{\mathscr{V}'})$ est un isomorphisme.

Par ailleurs, $\mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'}$ est l'idéal des éléments nilpotents de $\mathcal{O}_{\mathscr{V}'}$ (lemme 2.5) car $\mathscr{V}'$ est irréductible et génériquement étale sur $\tilde{\mathscr{S}}$. Donc on a la suite exacte de $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules :

$$0 \to \mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'} \to \mathcal{O}_{\mathscr{V}'} \to \mathcal{O}_{\tilde{\mathscr{V}}} \to 0.$$

Or $q_* \mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'} = 0$ car $q_* \mathcal{O}_{\mathscr{V}'} \simeq \mathcal{O}_{\mathscr{V}}$ est intègre.

En vertu de ce qui précède on en déduit $R^1 q_* \mathcal{O}_{\tilde{\mathscr{V}}} = 0$, et la suite exacte $0 \to \mathcal{O}_{\mathscr{V}} \to q_* \mathcal{O}_{\tilde{\mathscr{V}}} \to R^1 q_* \mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'} \to 0$.

L'algèbre $q_* \mathcal{O}_{\tilde{\mathscr{V}}}$ est intègre et $R^1 q_* \mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'}$ est à support ponctuel. Comme $\mathcal{O}_{\mathscr{V}}$ est normal, on a $R^1 q_* \mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}} \mathcal{O}_{\mathscr{V}'} = 0$.

Finalement, le $\mathcal{O}_{\tilde{\mathscr{S}}}$-module $\mathcal{O}_{\tilde{\mathscr{V}}}$ est localement libre si et seulement si $\mathrm{Ext}^1_{\mathcal{O}_{\tilde{\mathscr{S}}}}(\mathcal{O}_{\tilde{\mathscr{V}}}, \omega_{\tilde{\mathscr{S}}}) = 0$, où $\omega_{\tilde{\mathscr{S}}}$ est le faisceau dualisant de $\tilde{\mathscr{S}}$, et la dualité relative pour q donne un isomorphisme :

$$\mathrm{Ext}^1_{\mathcal{O}_{\tilde{\mathscr{S}}}}(\mathcal{O}_{\tilde{\mathscr{V}}}, \omega_{\tilde{\mathscr{S}}}) = \mathrm{Ext}^1_{\mathcal{O}_{\mathscr{S}}}(Rq_* \mathcal{O}_{\tilde{\mathscr{V}}}, K_{\mathscr{S}}[-2]),$$

où $K_{\mathscr{S}}$ est le complexe dualisant de $\mathscr{S}$. Mais on a $Rq_* \mathcal{O}_{\tilde{V}} \simeq \mathcal{O}_V$, d'où un isomorphisme :

$$\mathrm{Ext}^1_{\mathcal{O}_{\tilde{\mathscr{S}}}}(\mathcal{O}_{\tilde{\mathscr{V}}}, \omega_{\tilde{\mathscr{S}}}) \simeq \mathrm{Ext}^1_{\mathcal{O}_{\mathscr{S}}}(\mathcal{O}_{\mathscr{V}}, K_{\mathscr{S}}[-2]).$$

La dualité relative pour p donne un isomorphisme :

$$\mathrm{Ext}^1_{\mathcal{O}_{\mathscr{S}}}(\mathcal{O}_{\mathscr{V}}, K_{\mathscr{S}}) \simeq \mathrm{Ext}^1_{\mathcal{O}_{\mathscr{V}}}(\mathcal{O}_{\mathscr{V}}, K_{\mathscr{V}}[-2]),$$

mais comme $\mathscr{V}$ est normale, $\mathscr{V}$ est de Cohen-Macaulay, donc $K_{\mathscr{V}}[-2]$ n'a qu'un seul module de cohomologie non nul en degré 0. Par suite $\mathrm{Ext}^1_{\mathcal{O}_{\mathscr{V}}}(\mathcal{O}_{\mathscr{V}}, K_{\mathscr{V}}[-2]) = 0$, donc $\mathcal{O}_{\tilde{\mathscr{V}}}$ est localement libre, ce qui achève la démonstration de la proposition 2.4.

2.6. Remarque. — La proposition 2.4 montre que $\tilde{\mathscr{S}}$ est un platificateur de $\mathscr{V}$ sur $\mathscr{S}$. Nous verrons plus loin que, sous les hypothèses du théorème 2.2 (i. e. $\mathscr{V}$ lisse, $\mathscr{S}$ point double rationnel et $\tilde{\mathscr{S}}$ désingularisation minimale de $\mathscr{S}$) $\tilde{\mathscr{S}}$ est le platificateur normal de $\mathscr{V}$ sur $\mathscr{S}$.

2.7. L'application $\pi : R(G) \to K(\tilde{\mathscr{S}})$ du théorème 2.2 est $\mathbb{Z}$-linéaire, donc il suffit, pour déterminer π, de calculer $\pi(\rho)$ pour chaque représentation irréductible ρ de G.

Soit $\mathscr{M}_\rho$ le $\mathcal{O}_{\mathscr{S}}$-module défini par :

$$\mathscr{M}_\rho = \mathrm{Hom}_{C[G]}(E_\rho, \mathcal{O}_{\mathscr{V}}).$$

On pose $\mathscr{F}_\rho = (\mathcal{O}_{\tilde{\mathscr{S}}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathscr{M}_\rho)/\mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}}$, où $\mathrm{Tors}_{\mathcal{O}_{\tilde{\mathscr{S}}}}$ dénote la $\mathcal{O}_{\tilde{\mathscr{S}}}$-torsion de $\mathcal{O}_{\tilde{\mathscr{S}}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathscr{M}_\rho$.

Avec les notations précédentes on a :

2.8. PROPOSITION. — *Les $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules $(\mathcal{O}_{\tilde{\mathscr{V}}} \underset{\mathcal{O}_{\mathscr{V}}}{\otimes} (\mathcal{O}_{\mathscr{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee}))^{\mathrm{G}}$ et $\mathscr{F}_\rho$ sont isomorphes.*

Démonstration. — Notons $\mathscr{N}$ l'idéal des éléments nilpotents d'un anneau et T_{A} la torsion d'un A-module.

On a :

$$(\mathcal{O}_{\tilde{\mathscr{V}}} \underset{\mathcal{O}_{\mathscr{V}}}{\otimes} (\mathcal{O}_{\mathscr{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee}))^{\mathrm{G}}$$

$$\simeq (\mathcal{O}_{\tilde{\mathscr{V}}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})^{\mathrm{G}} \text{ (par simplification du produit tensoriel)}$$

$$= (((\mathcal{O}_{\mathscr{V}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathcal{O}_{\tilde{\mathscr{S}}})/\mathscr{N}) \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})^{\mathrm{G}} \text{ (définition de } \mathcal{O}_{\tilde{\mathscr{V}}})$$

$$= (((\mathcal{O}_{\mathscr{V}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathcal{O}_{\tilde{\mathscr{S}}})/\mathrm{T}_{\mathcal{O}_{\tilde{\mathscr{S}}}}) \underset{\mathbb{C}}{\otimes} \mathrm{E}_\rho)^{\mathrm{G}} \text{ (lemme 2.5)}$$

$$\simeq ((\mathcal{O}_{\mathscr{V}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathcal{O}_{\tilde{\mathscr{S}}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})/\mathrm{T}_{\mathcal{O}_{\tilde{\mathscr{S}}}})^{\mathrm{G}}$$

(commutativité de la torsion avec la somme directe)

$$\simeq (\mathcal{O}_{\mathscr{V}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} \mathcal{O}_{\tilde{\mathscr{S}}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})^{\mathrm{G}}/\mathrm{T}^{\mathrm{G}}_{\mathcal{O}_{\tilde{\mathscr{S}}}}$$

(commutativité de l'action de G avec la torsion)

$$\simeq (\mathcal{O}_{\tilde{\mathscr{S}}} \underset{\mathcal{O}_{\mathscr{S}}}{\otimes} (\mathcal{O}_{\mathscr{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})^{\mathrm{G}})/\mathrm{T}^{\mathrm{G}}_{\mathcal{O}_{\tilde{\mathscr{S}}}}$$

(car $\mathcal{O}_{\tilde{\mathscr{S}}}$ et $\mathcal{O}_{\mathscr{S}}$ sont invariants par G).

$\simeq \mathscr{F}_\rho$ car on a un isomorphisme canonique :

$$\mathscr{M}_\rho = \mathrm{Hom}_{\mathbb{C}[\mathrm{G}]}(\mathrm{E}_\rho, \mathcal{O}_{\mathscr{V}}) \xrightarrow{\sim} (\mathcal{O}_{\mathscr{V}} \underset{\mathbb{C}}{\otimes} \mathrm{E}_{\rho^\vee})^{\mathrm{G}}.$$

En vue de la définition de l'application π du théorème 2.2, on a :

2.9. COROLLAIRE (prop. 2.8). — *L'application $\pi : \mathrm{R}(\mathrm{G}) \to \mathrm{K}(\tilde{\mathscr{S}})$ est induite aussi par celle qui associe à chaque représentation irréductible ρ de G le $\mathcal{O}_{\tilde{\mathscr{S}}}$-module $\mathscr{F}_\rho$.*

2.10. COROLLAIRE. — (i) *On a $q_* \mathscr{F}_\rho = \mathscr{M}_\rho$, $\mathrm{R}^1 q_* \mathscr{F}_\rho = 0$.*

(ii) *Le $\mathcal{O}_{\tilde{\mathscr{S}}}$-module $\mathscr{F}_\rho$ est localement libre de rang égal au degré de ρ.*

Démonstration. — Notons $\mathring{\mathrm{Irr}}(\mathrm{G}) := \mathrm{Irr}(\mathrm{G}) \cup \{\rho_0\}$. Si A est un anneau, A [G] dénote la A-algèbre du groupe G.

Il résulte de la proposition 2.8 et de la décomposition canonique des G-modules qu'on a un isomorphisme canonique de $\mathcal{O}_{\mathcal{Z}}$ [G]-modules :

$$\mathcal{O}_{\tilde{\mathcal{V}}} \simeq \bigoplus_{\rho \in \mathring{\mathrm{Irr}}(\mathrm{G})} \mathcal{F}_\rho \underset{\mathbb{C}}{\otimes} \mathrm{E}_\rho$$

et par suite :

$$q_* \mathcal{O}_{\tilde{\mathcal{V}}} \simeq \bigoplus_{\rho \in \mathring{\mathrm{Irr}}(\mathrm{G})} q_* \mathcal{F}_\rho \underset{\mathbb{C}}{\otimes} \mathrm{E}_\rho,$$

$$\mathrm{R}^1 q_* \mathcal{O}_{\tilde{\mathcal{V}}} \simeq \bigoplus_{\rho \in \mathring{\mathrm{Irr}}(\mathrm{G})} \mathrm{R}^1 q_* \mathcal{F}_\rho \underset{\mathbb{C}}{\otimes} \mathrm{E}_\rho.$$

De même, on a un isomorphisme canonique de $\mathcal{O}_{\mathcal{S}}$ [G]-modules :

$$\mathcal{O}_{\mathcal{V}} \simeq \bigoplus_{\rho \in \mathring{\mathrm{Irr}}(\mathrm{G})} \mathcal{M}_\rho \underset{\mathbb{C}}{\otimes} \mathrm{E}_\rho$$

d'où (i) d'après la proposition 2.4 et le fait que $\mathcal{F}_\rho$, facteur direct d'un module localement libre, est localement libre.

Localement, en dehors du diviseur D, $\mathcal{O}_{\tilde{\mathcal{V}}}$ est isomorphe à $\mathcal{O}_{\mathcal{Z}}$ [G], et (ii) résulte alors de la structure de la représentation régulière.

3. Rappel sur les groupes engendrés par des réflexions; les modules M_ρ

3.1. Pour chaque sous-groupe fini G de SL(2, $\mathbb{C}$), il existe un produit scalaire hilbertien de $\mathrm{V} = \mathbb{C}^2$ invariant par G et dans la suite on supposera donc que G est un sous-groupe fini de SU(2, $\mathbb{C}$); on sait par ailleurs que deux sous-groupes de SU(2, $\mathbb{C}$) sont conjugués si et seulement si ils le sont dans SL(2, $\mathbb{C}$).

Pour la démonstration du théorème 2.2 on utilisera le fait que chaque groupe polyédral binaire G est un sous-groupe d'indice 2 d'un groupe $\mathrm{G}' \subset \mathrm{U}(2, \mathbb{C})$ engendré par des réflexions complexes d'ordre 2 (*voir* [S – T], [Co]). Ceci est dû essentiellement au fait que déjà les groupes polyédraux purs sont des sous-groupes d'indice 2 de groupes engendrés par des réflexions de $\mathbb{R}^3$; ces derniers sont obtenus en considérant, en plus des rotations, les réflexions qui laissent fixe un polyèdre (*voir* [K], chap. 1, § 11 ou [D], § 12). Nous utiliserons le fait que G′/G opère sur les représentations de G en associant à chaque représentation ρ, sa contragrédiente $\check{\rho}$.

3.2. Notons $\mathrm{S}' = \mathrm{V}/\mathrm{G}'$ la surface obtenue comme quotient de V par G′, A_V l'anneau de polynômes à deux variables à coefficients dans $\mathbb{C}$, A_S (resp. $\mathrm{A}_{\mathrm{S}'}$) l'anneau de fonctions régulières globales sur S (resp. sur S′). Dans la suite on identifie l'anneau A_S (resp. $\mathrm{A}_{\mathrm{S}'}$) au sous-anneau de A_V des invariants de A_V par G (resp. G′).

On sait que A_V est un $A_{S'}$-module libre, et que $A_{S'}$ est une $\mathbb{C}$-algèbre (graduée) de polynômes ([Ch], [Bo], chap. V, § 5).

Notons Y, Z deux invariants homogènes de G′ qui engendrent $A_{S'}$ comme $\mathbb{C}$-algèbre; et considérons $X = J(Y, Z)$ le déterminant jacobien de Y et Z.

Alors X est un semi-invariant de G′ (*cf.* [Bo]), dont le carré est invariant par G′ car les réflexions de G′ sont d'ordre 2; et X, Y, Z sont trois invariants homogènes de G qui engendrent A_S comme $\mathbb{C}$-algèbre ([K], chap. 2).

Pour fixer les notations, soient Y et Z choisis de telle façon que $\deg(Z) \leqq \deg(Y)$; cette convention détermine X à un facteur constant près (sauf si G est de type $\mathbf{A}_1$), de même pour Y (sauf si G est de type $\mathbf{A}_n$ avec n impair ou $\mathbf{D}_n$ avec n pair), de même pour Z (sauf si G est de type $\mathbf{A}_1$ ou $\mathbf{D}_4$).

Dans le tableau suivant on donne les degrés de X, Y, Z dans chaque cas :

Type de G	deg(X)	deg(Y)	deg(Z)
$\mathbf{A}_n (n \geqq 1)$	$n+1$	$n+1$	2
$\mathbf{D}_n (n \geqq 4)$	$2(n-1)$	$2(n-2)$	4
$\mathbf{E}_6$	12	8	6
$\mathbf{E}_7$	18	12	8
$\mathbf{E}_8$	30	20	12

3.3. Notons ε la représentation de degré 1 de G′ définie par l'action de G′ sur $\mathbb{C}X$, où X est le semi-invariant considéré au paragraphe précédent.

Soit ρ (resp. ρ') une représentation de G (resp. G′). Par abus de notation on écrira $\varepsilon\rho'$ au lieu de $\varepsilon \otimes \rho$, $\operatorname{Ind}\rho$ au lieu de $\operatorname{Ind}_G^{G'}\rho$ (i. e. la représentation induite sur G′ par ρ [Se], § 7.1) et $\operatorname{Res}\rho'$ au lieu de $\operatorname{Res}_G \rho'$ (i. e. la représentation restreinte à G de ρ').

3.4. PROPOSITION. — *Soit* ρ *une représentation irréductible de* G. *Alors on a :*

(i) $\varepsilon \operatorname{Ind}\rho = \operatorname{Ind}\rho = \operatorname{Ind}\rho^{\vee}$.

(ii) $\rho \neq \rho^{\vee}$ *si et seulement si* $\operatorname{Ind}\rho$ *est irréductible.*

(iii) $\rho = \rho^{\vee}$ *si et seulement si* $\operatorname{Ind}\rho = \rho' \oplus \varepsilon\rho'$, *où* ρ' *est une représentation irréductible de* G′.

Soit ρ' *une représentation irréductible de* G′. *Alors on a :*

(i′) $\operatorname{Res}\rho' = \operatorname{Res}\varepsilon\rho' = (\operatorname{Res}\rho')^{\vee}$.

(ii′) $\rho' \neq \varepsilon\rho'$ *si et seulement si* $\operatorname{Res}\rho'$ *est irréductible.*

(iii′) $\rho' = \varepsilon\rho'$ *si et seulement si* $\operatorname{Res}\rho' = \rho \oplus \rho^{\vee}$, *où* ρ *est une représentation irréductible de* G.

Démonstration. — Quelque soit ρ' représentation de G′ on a :

$$(\operatorname{Ind}\rho) \otimes \rho' = \operatorname{Ind}(\rho \otimes \operatorname{Res}\rho'), \tag{1}$$

donc, si $\rho' = \varepsilon$ on a :

$$\varepsilon \operatorname{Ind}\rho = \operatorname{Ind}\rho, \tag{2}$$

car $\operatorname{Res}\varepsilon = \rho_0$, où ρ_0 est la représentation triviale de rang 1 de G; et si $\rho = \rho_0$ dans (1), alors :

$$\operatorname{Ind}\operatorname{Res}\rho' = \rho' \oplus \varepsilon\rho', \tag{3}$$

car $\operatorname{Ind}\rho_0=\rho'_0\oplus\varepsilon$, où ρ'_0 est la représentation triviale de rang 1 de G'. Par ailleurs, le groupe quotient G'/G agit sur les représentations de G en associant à chaque représentation ρ la contragrédiente $\rho^{\vee}$. Par suite on a :

$$\operatorname{Res}\operatorname{Ind}\rho=\rho\oplus\rho^{\vee}. \tag{4}$$

En appliquant Ind aux deux membres de (4), on a, par (3) :

$$\operatorname{Ind}\rho\oplus\varepsilon\operatorname{Ind}\rho=\operatorname{Ind}\rho\oplus\operatorname{Ind}\rho^{\vee},$$

par suite $\varepsilon\operatorname{Ind}\rho=\operatorname{Ind}\rho^{\vee}$, d'où (i), par (2).

Par le critère d'irréductibilité de Mackey ([Se], § 7.5), on a que $\operatorname{Ind}\rho$ est irréductible si et seulement si $\rho\neq\rho^{\vee}$, d'où (ii), et on a $\rho=\rho^{\vee}$ si et seulement si $\operatorname{Ind}\rho=\bigoplus_{i=1}^{m}\rho'_i$, avec $m>1$ et ρ'_i irréductible pour tout i. Or en appliquant Res aux deux membres on a , par (4), une somme de deux représentations irréductibles, donc $m\leqq 2$ et par suite $m=2$. Si $\rho=\rho^{\vee}$ on a donc

$$\operatorname{Res}(\rho'_1\oplus\rho'_2)=\rho\oplus\rho,$$

par suite

$$\operatorname{Res}\rho'_1=\operatorname{Res}\rho'_2=\rho,\quad \text{d'où}\quad \rho'_1\oplus\rho'_2=\operatorname{Ind}\rho=\operatorname{Ind}\operatorname{Res}\rho'_1=\rho'_1\oplus\varepsilon\rho'_1,$$

la dernière égalité par (3), d'où $\varepsilon\rho'_1=\rho'_2$, ce qui montre (iii).

Par ailleurs on a $\operatorname{Res}\varepsilon\rho'=(\operatorname{Res}\varepsilon)\otimes(\operatorname{Res}\rho')=\operatorname{Res}\rho'$, car $\operatorname{Res}\varepsilon=\rho_0$; et par (4) on a :

$$(\operatorname{Res}\rho')\oplus(\operatorname{Res}\rho')^{\vee}=\operatorname{Res}\operatorname{Ind}\operatorname{Res}\rho'=\operatorname{Res}(\rho'\oplus\varepsilon\rho')=(\operatorname{Res}\rho')\oplus(\operatorname{Res}\rho')$$

par suite $\operatorname{Res}\rho'=(\operatorname{Res}\rho')^{\vee}$, d'où (i').

Soit $\operatorname{Res}\rho'=\bigoplus_{i=1}^{m}\rho_i$ la décomposition en somme de représentation irréductibles de $\operatorname{Res}\rho'$; alors, par (3), on a $m\leqq 2$.

Si $\operatorname{Res}\rho'=\rho_1\oplus\rho_2$, alors $\rho'\oplus\varepsilon\rho'=\operatorname{Ind}\operatorname{Res}\rho'=(\operatorname{Ind}\rho_1)\oplus(\operatorname{Ind}\rho_2)$ et par suite $\operatorname{Ind}\rho_i$ est irréductible ($i=1$, 2), et on peut supposer $\operatorname{Ind}\rho_1=\rho'$, $\operatorname{Ind}\rho_2=\varepsilon\rho'$. En ce cas on a, par (4), $\rho_1\oplus\rho_1^{\vee}=\operatorname{Res}\operatorname{Ind}\rho_1=\operatorname{Res}\rho'=\rho_1\oplus\rho_2$, d'où $\rho_1^{\vee}=\rho_2$ et par (i) on a $\rho'=\varepsilon\rho'$.

Finalement, si $\operatorname{Res}\rho'=\rho$, alors par (3) on a $\operatorname{Ind}\rho=\rho'\oplus\varepsilon\rho'$. Or par la formule de réciprocité de Frobenius on a :

$$1=(\chi_\rho\mid\chi_{\operatorname{Res}\rho'})_G=(\chi_{\operatorname{Ind}\rho}\mid\chi_{\rho'})_{G'},$$

où χ dénote le caractère, et par conséquent $\rho'\neq\varepsilon\rho'$, ce qui achève la démonstration.

3.5. Soit ρ (resp. ρ') une représentation irréductible de G (resp. de G'). On considère le A_S-module gradué (resp. le $A_{S'}$-module gradué) M_ρ (resp. $M_{\rho'}$) défini par :

$$M_\rho=\operatorname{Hom}_{C[G]}(E_\rho, A_V)\quad [\text{resp. } M_{\rho'}=\operatorname{Hom}_{C[G']}(E_{\rho'}, A_V)].$$

On pose $U_\rho = M_\rho / \mathscr{M}_{S'} M_\rho$ et $U_{\rho'} = M_{\rho'} / \mathscr{M}_{S'} M_{\rho'}$, où $\mathscr{M}_{S'}$ est l'idéal maximal de $A_{S'}$ des éléments de degré strictement positif [i. e. $\mathscr{M}_{S'} = (Y, Z)$]; U_ρ (resp. $U_{\rho'}$) est un $\mathbb{C}$-espace vectoriel muni de la graduation canonique induite par celle de M_ρ (resp. $M_{\rho'}$). Il résulte de la proposition 3.4 le corollaire suivant :

3.6. Corollaire. — (i) *Si* $\rho \neq \check{\rho}$, *soit* $\rho' = \operatorname{Ind} \rho$. *On a des* $A_{S'}$*-isomorphismes* (*resp. des* $\mathbb{C}$*-isomorphismes*) *homogènes de degré* 0 :

$$M_\rho = M_{\rho'} = M_{\check{\rho}} \quad (\textit{resp. } U_\rho = U_{\rho'} = U_{\check{\rho}})$$

(ii) *Si* $\rho = \check{\rho}$, *soit* $\operatorname{Ind} \rho = \rho' \oplus \varepsilon \rho'$ *la décomposition en représentations irréductibles de* $\operatorname{Ind} \rho$. *On a un* $A_{S'}$*-isomorphisme* (*resp. un* $\mathbb{C}$*-isomorphisme*) *homogène de degré* 0 :

$$M_\rho = M_{\rho'} \oplus M_{\varepsilon\rho'} \quad (\textit{resp. } U_\rho = U_{\rho'} \oplus U_{\varepsilon\rho'}).$$

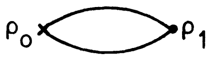
3.7. Notons c' la représentation canonique de G' définie par l'inclusion $G' \subset U(2, \mathbb{C})$. On peut définir un diagramme $\tilde{\Gamma}(G')$, de façon analogue à celle qui définit $\tilde{\Gamma}(G)$, en considérant maintenant le produit tensoriel par la représentation c'. Or, étant donné qu'on a en général $c' \neq \varepsilon c'$ (en fait on a égalité seulement si G est de type $\mathbf{A}_n$), on introduit lorsque $c' \neq \varepsilon c'$ des arêtes orientées : si ρ'_1 et ρ'_2 sont deux représentations irréductibles de G', on joint les deux sommets correspondants par une arête orientée de source ρ'_1 et de but ρ'_2 si ρ'_2 apparaît dans la décomposition de $c' \otimes \rho'_1$ en représentations irréductibles.

Dans la liste suivante on donne les diagrammes qu'on obtient. On utilise les notations de la proposition 3.4, et on fixe une numérotation des sommets qui sera maintenue dans la suite. Le sommet correspondant à la représentation triviale est indiqué par une croix.

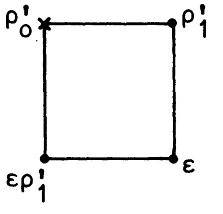
3.7.1. *Type* $\mathbf{A}_n$ $(n \geqq 1)$

(*a*) $n = 1$:

$\tilde{\mathbf{A}}_1$:
$(c = 2\rho_1)$

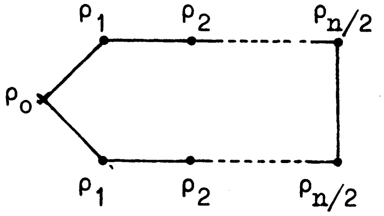


$\tilde{\mathbf{A}}'_1$:
$(c' = \rho'_1 \oplus \varepsilon\rho'_1)$

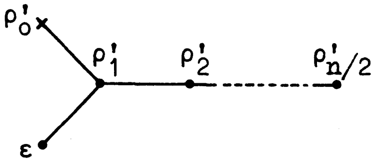


(*b*) *n* pair :

$\tilde{\mathbf{A}}_n$:
$(c = \rho_1 \oplus \rho_1)$

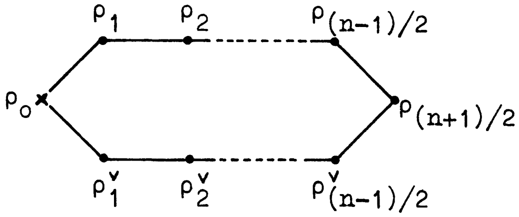


$\tilde{\mathbf{A}}'_n$:
$(c'=\rho'_1)$

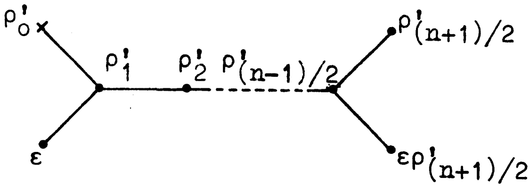

ρ'0
ρ'1
ρ'2
ρ'n/2
ε

(*c*) *n* impair, $n>1$:

$\tilde{\mathbf{A}}_n$:
$(c=\rho_1\oplus\check{\rho}_1)$


ρ1
ρ2
ρ(n−1)/2
ρ0
ρ(n+1)/2
ρ̌1
ρ̌2
ρ̌(n−1)/2

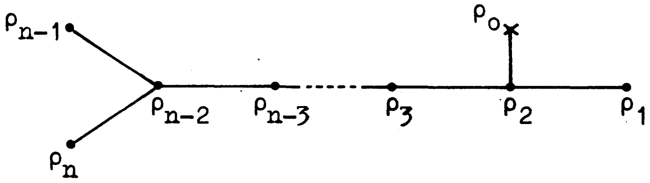
$\tilde{\mathbf{A}}'_n$:
$(c'=\rho'_1)$


ρ'0
ρ'1
ρ'2
ρ'(n−1)/2
ρ'(n+1)/2
ε
ερ'(n+1)/2

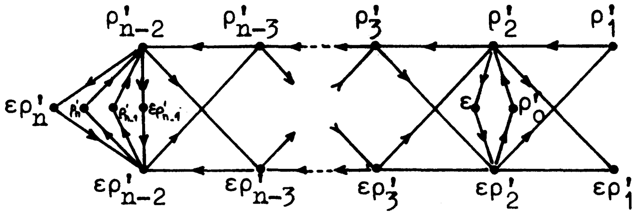
3.7.2. *Type* $\mathbf{D}_n$ $(n\geqq 4)$

(*a*) *n* pair :

$\tilde{\mathbf{D}}_n$:
$(c=\rho_2)$

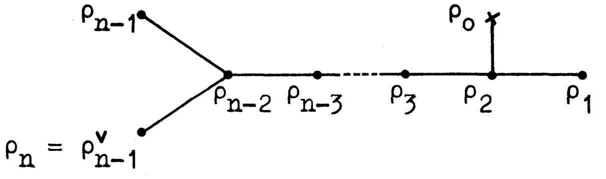

ρn−1
ρ0
ρn−2
ρn−3
ρ3
ρ2
ρ1
ρn

$\tilde{\mathbf{D}}'_n$:
$(c'=\rho'_2)$

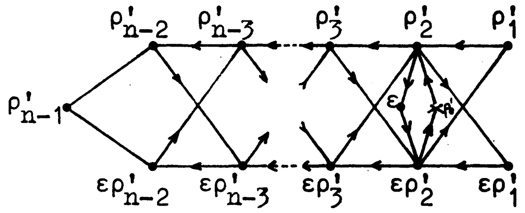

ρ'n−2
ρ'n−3
ρ'3
ρ'2
ρ'1
ερ'n
ρ'n
ρ'n−1
ερ'n−1
ε
ρ'0
ερ'n−2
ερ'n−3
ερ'3
ερ'2
ερ'1

(*b*) *n* impair :

$\tilde{\mathbf{D}}_n$:
$(c=\rho_2)$

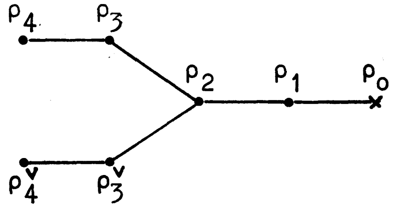

ρn−1
ρ0
ρn−2
ρn−3
ρ3
ρ2
ρ1
ρn = ρ̌n−1

$\tilde{\mathbf{D}}'_n$:
$(c=\rho'_2)$

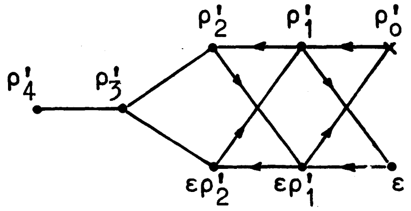

ρ'_{n-2}
ρ'_{n-3}
ρ'_3
ρ'_2
ρ'_1
ρ'_{n-1}
ε
ρ'_0
$\varepsilon\rho'_{n-2}$
$\varepsilon\rho'_{n-3}$
$\varepsilon\rho'_3$
$\varepsilon\rho'_2$
$\varepsilon\rho'_1$

3.7.3. *Type* $\mathbf{E}_6$

$\tilde{\mathbf{E}}_6$:
$(c=\rho_1)$

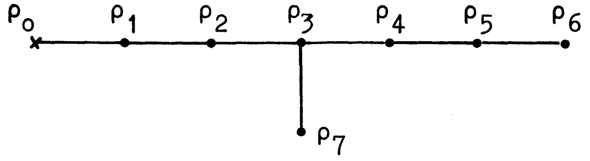

ρ_4
ρ_3
ρ_2
ρ_1
ρ_0
$\rho_4^{\vee}$
$\rho_3^{\vee}$

$\tilde{\mathbf{E}}'_6$:
$(c'=\rho'_1)$

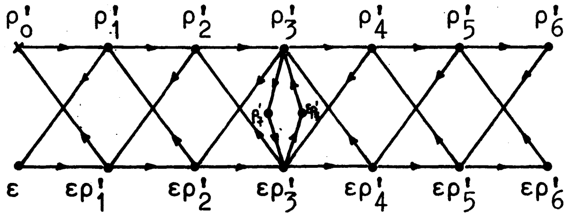

ρ'_2
ρ'_1
ρ'_0
ρ'_4
ρ'_3
$\varepsilon\rho'_2$
$\varepsilon\rho'_1$
ε

3.7.4. *Type* $\mathbf{E}_7$

$\tilde{\mathbf{E}}_7$:
$(c=\rho_1)$

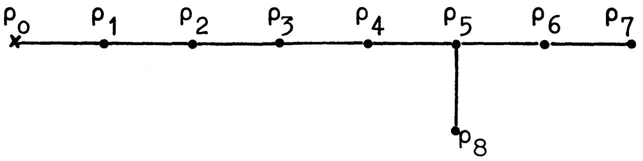

ρ_0
ρ_1
ρ_2
ρ_3
ρ_4
ρ_5
ρ_6
ρ_7

$\tilde{\mathbf{E}}'_7$:
$(c'=\rho'_1)$

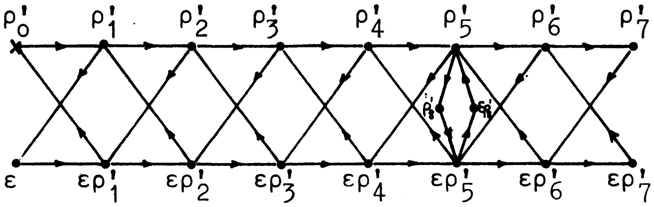

ρ'_0
ρ'_1
ρ'_2
ρ'_3
ρ'_4
ρ'_5
ρ'_6
ρ'_7
$\varepsilon\rho'_7$
ε
$\varepsilon\rho'_1$
$\varepsilon\rho'_2$
$\varepsilon\rho'_3$
$\varepsilon\rho'_4$
$\varepsilon\rho'_5$
$\varepsilon\rho'_6$

3.7.5. *Type* $\mathbf{E}_8$

$\tilde{\mathbf{E}}_8$:
$(c=\rho_1)$


ρ_0
ρ_1
ρ_2
ρ_3
ρ_4
ρ_5
ρ_6
ρ_7
ρ_8

$\tilde{\mathrm{E}}'_8$:
$(c'=\rho'_1)$



4. Résolution G-équivariante de $\mathrm{A_V}$; l'opérateur Θ

4.1. On utilise les notations du paragraphe précédent. L'anneau $\mathrm{A_V}$ est muni d'une structure de $\mathrm{A_S}\,[\mathrm{G}]$-module gradué (resp. de $\mathrm{A_{S'}}\,[\mathrm{G'}]$-module gradué, resp. de $\mathrm{A_{S'}}\,[\mathrm{G}]$-module gradué).

Soit $\Theta : \mathrm{A_V} \to \mathrm{A_V}$ l'opérateur $\mathrm{A_{S'}}\,[\mathrm{G}]$-linéaire défini par $\Theta f = \mathrm{X} f$ pour $f \in \mathrm{A_V}$; il est homogène de degré égal au degré de X et il est aussi $\mathrm{A_S}\,[\mathrm{G}]$-linéaire.

Notons $\sigma : \mathrm{A_S} \to \mathrm{A_S}$ l'automorphisme d'ordre 2 de $\mathrm{A_{S'}}$-algèbre tel que $\sigma(\mathrm{X}) = -\mathrm{X}$. Pour tout $f \in \mathrm{A_V}$ et $h \in \mathrm{A_S}$ posons $h\,._{\sigma} f = \sigma(h) f$. On obtient ainsi une nouvelle structure de $\mathrm{A_S}$-module sur $\mathrm{A_V}$; cet $\mathrm{A_S}$-module est noté $\mathrm{A_V^{\sigma}}$.

Considérons le $\mathrm{A_{S'}}\,[\mathrm{G'}]$-module $\mathrm{A_V} \underset{\mathrm{A_{S'}}}{\otimes} \mathrm{A_S}$; il est aussi muni d'une structure de $\mathrm{A_S}\,[\mathrm{G}]$-module, et il est libre en tant que $\mathrm{A_S}$-module.

4.2. Lemme. — (i) *Le rang de* $\mathrm{A_V} \underset{\mathrm{A_{S'}}}{\otimes} \mathrm{A_S}$ *comme* $\mathrm{A_{S'}}$*-module est égal à* $2\,|\,\mathrm{G'}\,|$ *— où* $|\,\mathrm{G'}\,|$ *dénote l'ordre de* $\mathrm{G'}$ *—; et son rang comme* $\mathrm{A_S}$*-module est* $2\,|\,\mathrm{G}\,|$.

(ii) *On a deux suites exactes de* $\mathrm{A_S}\,[\mathrm{G}]$*-modules :*

$$0 \to \mathrm{A_V^{\sigma}} \xrightarrow{i^-} \mathrm{A_V} \underset{\mathrm{A_{S'}}}{\otimes} \mathrm{A_S} \xrightarrow{j^+} \mathrm{A_V} \to 0,$$

$$0 \to \mathrm{A_V} \xrightarrow{i^+} \mathrm{A_V} \underset{\mathrm{A_{S'}}}{\otimes} \mathrm{A_S} \xrightarrow{j^-} \mathrm{A_V^{\sigma}} \to 0$$

avec $i^+(f) = i^-(f) = f \otimes \mathrm{X} - \Theta f \otimes 1$ *:*

$$j^+(f \otimes h) = fh, \qquad j^-(f \otimes h) = h\,._{\sigma} f,$$
$$\forall f \in \mathrm{A_V}, \qquad \forall h \in \mathrm{A_S}.$$

Démonstration. — On a un isomorphisme de $\mathrm{A_{S'}}\,[\mathrm{G'}]$-modules $\mathrm{A_V} \simeq \mathrm{A_{S'}} \underset{\mathbb{C}}{\otimes} \mathrm{A_V} / \mathscr{M}_{\mathrm{S'}} \mathrm{A_V}$, où $\mathscr{M}_{\mathrm{S'}}$ dénote l'idéal maximal de $\mathrm{A_{S'}}$ engendré par Y et Z [Bo]; et d'autre part le $\mathbb{C}\,[\mathrm{G}]$-module gradué $\mathrm{A_V} / \mathscr{M}_{\mathrm{S'}} \mathrm{A_V}$ fournit une représentation régulière de $\mathrm{G'}$, d'où (i). La vérification de l'exactitude des suites de (ii) est immédiate.

4.3. Considérons par ailleurs l'opérateur $\mathrm{A_{S'}}\,[\mathrm{G}]$-linéaire homogène de $\mathrm{A_S}$ dans lui-même défini par la multiplication par X; on le notera aussi X. Il est aussi $\mathrm{A_S}\,[\mathrm{G}]$-linéaire.

4.4. PROPOSITION. — (i) *On a l'égalité* $\Theta^2 = X^2 \, \mathrm{Id}_{A_V}$ *où* $X^2 \, \mathrm{Id}_{A_V}$ *désigne la multiplication par* $X^2 \in A_{S'}$.

(ii) *Soient* T^+ (*resp.* T^-) : $A_V \otimes_{A_{S'}} A_S \to A_V \otimes_{A_{S'}} A_S$ *l'opérateur* $A_S[G]$-*linéaire homogène défini par :*

$$T^+ = \Theta \otimes \mathrm{Id}_{A_S} + \mathrm{Id}_{A_V} \otimes X \quad (\textit{resp.}\ T^- = \Theta \otimes \mathrm{Id}_{A_S} - \mathrm{Id}_{A_V} \otimes X).$$

Alors on a la suite exacte de $A_S[G]$-*modules :*

$$\ldots \xrightarrow{T^-} A_V \otimes_{A_{S'}} A_S \xrightarrow{T^+} A_V \otimes_{A_{S'}} A_S \xrightarrow{T^-} A_V \otimes_{A_{S'}} A_S \xrightarrow{j^+} A_V \to 0$$

où j^+ *est défini dans le lemme* 4.2. *On obtient ainsi une résolution du* $A_S[G]$-*module gradué* A_V *par des* $A_S[G]$-*modules gradués libres en tant que* A_S-*modules.*

Démonstration. — L'opérateur Θ est défini comme étant la multiplication par X, or $X^2 \in A_{S'}$, d'où (i). D'autre part on a $T^+ = i^+ \circ j^+$ et $T^- = i^- \circ j^-$. L'exactitude de la suite de (ii) est donc conséquence du lemme 4.1 (ii).

4.5. Pour chaque représentation irréductible ρ de G, on considère le module M_ρ comme $A_{S'}$-module (§ 3.5).

Soit $\theta_\rho = M_\rho \to M_\rho$ (resp. $j_\rho^+ : M_\rho \otimes_{A_{S'}} A_S \to M_\rho$) l'application $A_{S'}$-linéaire obtenue en appliquant $\mathrm{Hom}_{\mathbb{C}[G]}(E_\rho, -)$ à Θ (resp. à j^+). Posons :

$$T_\rho^+ = \theta_\rho \otimes \mathrm{Id}_{A_S} + \mathrm{Id}_{M_\rho} \otimes X$$
$$T_\rho^- = \theta_\rho \otimes \mathrm{Id}_{A_S} - \mathrm{Id}_{M_\rho} \otimes X.$$

On définit ainsi des opérateurs $A_{S'}$-linéaires (resp. A_S-linéaires) homogènes de $M_\rho \otimes_{A_{S'}} A_S$ dans lui-même. On a alors :

4.6. COROLLAIRE. — *On a la suite exacte de* A_S-*modules gradués :*

$$\ldots \xrightarrow{T_\rho^-} M_\rho \otimes_{A_{S'}} A_S \xrightarrow{T_\rho^+} M_\rho \otimes_{A_{S'}} A_S \xrightarrow{T_\rho^-} M_\rho \otimes_{A_{S'}} A_S \xrightarrow{j_\rho^+} M_\rho \to 0.$$

Cette suite exacte fournit une résolution de M_ρ *par des* A_S-*modules gradués libres* (*sur* A_S) *de rang* 2 deg ρ.

4.7. REMARQUE. — Pour $\rho \neq \rho_0$, on peut montrer que cette résolution est une résolution minimale.

5. La série de Poincaré; graduations et matrices

5.1. Soit $\rho \in \mathring{\mathrm{I}}\mathrm{rr}(G)$ [resp. $\rho' \in \mathring{\mathrm{I}}\mathrm{rr}(G')$] une représentation irréductible de G (resp. de G'). On considère le $\mathbb{C}$-espace vectoriel gradué $U_\rho = M_\rho / \mathscr{M}_{S'} M_\rho$ (resp. $U_{\rho'} = M_{\rho'} / \mathscr{M}_{S'} M_{\rho'}$) de dimension égal à 2 deg ρ (resp. égale à deg ρ'). Notons $U_{\rho,n}$ (resp. $U_{\rho',n}$) le sous-espace des éléments homogènes de degré n.

Posons :

$$P_\rho(T)=\sum_n(\dim_{\mathbb{C}} U_{\rho,n})\,T^n, \qquad P_{\rho'}(T)=\sum_n(\dim_{\mathbb{C}} U_{\rho',n})\,T^n;$$

ce sont des polynômes dans $\mathbb{Z}(T)$.

Soit $R(G)[T]$ (resp. $R(G')[T]$) l'anneau des polynômes à une variable à coefficients dans $R(G)$ [resp. dans $R(G')$].

Soit c (resp. c') la représentation canonique de G (resp. de G') et ρ_0 (resp. ρ'_0) la représentation triviale de degré 1 de G (resp. de G').

5.2. Proposition. — *Dans l'anneau* $R(G)[T]$ (*resp.* $R(G')[T]$) *on a l'égalité :*

$$\Big(\sum_{\rho\in \operatorname{Irr}(G)} \rho\, P_\rho(T)\Big)(\rho_0(1+T^2)-c\,T)=\rho_0(1-T^{\deg Y})(1-T^{\deg Z})$$

$$\Big[\textit{resp. } \Big(\sum_{\rho'\in \operatorname{Irr}(G')} \rho'\, P_{\rho'}(T)\Big)(\rho'_0-c'\,T+\varepsilon T^2)=\rho'_0(1-T^{\deg Y})(1-T^{\deg Z})\Big].$$

Démonstration. — Nous ferons la démonstration pour G, celle pour G' est analogue. Si M est un $\mathbb{C}[G]$-module, on note $[M]$ sa classe dans $R(G)$. Soit $A_V=\bigoplus_{n=0}^{\infty} A_{V,n}$ la décomposition en sous-espaces homogènes de A_V. On définit la série de Poincaré $\langle A_V\rangle$ dans l'anneau de séries formelles $R(G)[[T]]$ par :

$$\langle A_V\rangle=\sum_{n=0}^{\infty}[A_{V,n}]\,T^n.$$

Étant donnée que $[A_{V,n}]=[\operatorname{sym}^n(c)]$, alors on a :

$$\langle A_V\rangle=\sum_{n=0}^{\infty}[\operatorname{sym}^n(c)]\,T^n=1\Big/\Big(\sum_{n=0}^{\infty}(-1)^n[\Lambda^n(c)]\,T^n\Big).$$

Or $\Lambda^0(c)=\rho_0$, $\Lambda^1(c)=c$, $\Lambda^2(c)=\rho_0$ car $G\subset SU(2,\mathbb{C})$, et $\Lambda^n(c)=0$, $\forall n>2$. Par suite on a :

$$(1)\qquad \langle A_V\rangle=1/(\rho_0(1+T^2)-c\,T).$$

Considérons par ailleurs la décomposition de A_V en $\mathbb{C}[G]$-sous-modules isotypiques, $A_V=\bigoplus_{\rho\in\operatorname{Irr}(G)} A_{V,\rho}$, et soit $\langle A_{V,\rho}\rangle$ la série de Poincaré de $A_{V,\rho}$ dans $R(G)[[T]]$. Alors on a :

$$(2)\qquad \langle A_V\rangle=\sum_{\rho\in\operatorname{Irr}(G)}\langle A_{V,\rho}\rangle.$$

Posons $S_\rho(T)=\sum_{n=0}^{\infty}(\dim_{\mathbb{C}} M_{\rho,n})\,T^n\in\mathbb{C}[[T]]$. Par la définition de M_ρ on a un isomorphisme homogène $A_{V,\rho}=M_\rho \underset{\mathbb{C}}{\otimes} E_\rho$. Par suite on a :

$$\langle A_{V,\rho}\rangle=\rho\, S_\rho(T). \tag{3}$$

Finalement on a $M_\rho=U_\rho \underset{\mathbb{C}}{\otimes} A_{S'}$, d'où :

$$S_\rho(T)=P_\rho(T)/(1-T^{\deg Y})(1-T^{\deg Z}). \tag{4}$$

La proposition résulte des égalités (1) à (4).

5.3. La proposition précédente fournit une méthode effective pour le calcul de la graduation de U_ρ (resp. de $U_{\rho'}$), qui est donné par le polynôme $P_\rho(T)$ [resp. $P_{\rho'}(T)$]. Par ailleurs, on a (cor. 3.6) :

$$\begin{aligned} P_\rho(T)=P_{\check{\rho}}(T)=P_{\rho'}(T) \qquad &\text{si} \quad \rho\neq\check{\rho} \quad \text{et} \quad \rho'=\operatorname{Ind}\rho,\\ P_\rho(T)=P_{\rho'}(T)+P_{\varepsilon\rho'}(T) \qquad &\text{si} \quad \rho=\check{\rho} \quad \text{et} \quad \operatorname{Ind}\rho=\rho'\oplus\varepsilon\rho'; \end{aligned}$$

par conséquent, il sugfit de calculer $P_{\rho'}(T)$ pour toute représentation irréductible ρ' de G' pour déterminer $P_\rho(T)$ pour toute représentation irréductible ρ de G.

Dans la table suivante on donne la liste des résultats. Le calcul est réalisé en utilisant le fait que $P_{\rho'_0}(T)=1$ et $P_\varepsilon(T)=T^{\deg X}$. Si $P_{\rho'}(T)=\sum_{i=1}^{k} a_i T^{n_i}$, on donne la suite des degrés n_i, avec chaque entier n_i répété a_i fois, $i=1,\ldots,k$.

Les notations utilisées pour les représentations sont celles introduites au paragraphe 3.7.

5.3.1. *Type* $\mathbf{A}_n (n\geqq 1)$

	représentation	*degrés*
(*a*) *n impair :*	$\rho'_j (1\leqq j\leqq (n-1)/2)$	$j; (n+1-j)$
	$\rho'_{(n+1)/2}$	$(n+1)/2$
	$\varepsilon\rho'_{(n+1)/2}$	$(n+1)/2$
(*b*) *n pair :*	$\rho'_j (1\leqq j\leqq n/2)$	$j; (n+1-j)$

5.3.2. *Type* $\mathbf{D}_n (n\geqq 4)$

(*a*) *n pair :*	ρ'_1	$(2n-4)$
	$\varepsilon\rho'_1$	2
$2\leqq j\leqq n-2$	ρ'_j (j pair)	$(j-1); (2n-j-1)$
	$\varepsilon\rho'_j$ (j pair)	$(j+1); (2n-j-3)$
	ρ'_j (j impair)	$(j-1); (2n-j-3)$
	$\varepsilon\rho'_j$ (j impair)	$(j+1); (2n-j-1)$
	$\rho'_j (j=n-1, n)$	n
	$\varepsilon\rho'_j (j=n-1, n)$	$n-2$

(*b*) *n impair* : Pour $j\leqq n-2$, mêmes résultats que ceux du cas n pair.

ρ'_{n-1}	$n-2; n$

5.3.3. *Type* $\mathbf{E}_6$

représentation	*degrés*
ρ'_1	1; 11
$\varepsilon\rho'_1$	5; 7
ρ'_2	2; 4; 6
$\varepsilon\rho'_2$	6; 8; 10
ρ'_3	3; 5; 7; 9
ρ'_4	4; 8

5.3.4. *Type* $\mathbf{E}_7$

représentation	*degrés*	*représentation*	*degrés*
ρ'_1	1; 17	$\varepsilon\rho'_1$	7; 11
ρ'_2	2; 6; 10	$\varepsilon\rho'_2$	8; 12; 16
ρ'_3	3; 7; 11; 15	$\varepsilon\rho'_3$	5; 9; 9; 13
ρ'_4	4; 8; 12	$\varepsilon\rho'_4$	6; 10; 14
ρ'_5	5; 13	$\varepsilon\rho'_5$	7; 11
ρ'_6	6	$\varepsilon\rho'_6$	12
ρ'_7	4; 8	$\varepsilon\rho'_7$	10; 14

5.3.5. *Type* $\mathbf{E}_8$

représentation	*degrés*	*représentation*	*degrés*
ρ'_1	1; 29	$\varepsilon\rho'_1$	11; 19
ρ'_2	2; 10; 18	$\varepsilon\rho'_2$	12; 20; 28
ρ'_3	3; 11; 19; 27	$\varepsilon\rho'_3$	9; 13; 17; 21
ρ'_4	4; 8; 12; 16; 20	$\varepsilon\rho'_4$	10; 14; 18; 22; 26
ρ'_5	5; 9; 13; 17; 21; 25	$\varepsilon\rho'_5$	7; 11; 15; 15; 19; 23
ρ'_6	6; 14; 18; 22	$\varepsilon\rho'_6$	8; 12; 16; 24
ρ'_7	7; 23	$\varepsilon\rho'_7$	13; 17
ρ'_8	6; 10; 14	$\varepsilon\rho'_8$	16; 20; 24

5.4. DÉFINITION DES MATRICES $[\theta_\rho]$; QUELQUES REMARQUES UTILES.

5.4.1. Soit M un $\mathrm{A}_{\mathrm{S}'}$-module gradué libre de type fini. Posons $\mathrm{U}=\mathrm{M}/\mathscr{M}_{\mathrm{S}'}\mathrm{M}$; U est gradué canoniquement par la graduation induite par celle de M. Notons M_n (resp. U_n) le sous-groupe de M (resp. de U) des éléments homogènes de degré n. Soit n un entier tel que $\mathrm{U}_n \neq 0$. On considère le sous-ensemble I_n de $\mathbb{Z}$ défini par :

$$\mathrm{I}_n=\{m \mid m=n+i, \text{ avec } \mathrm{U}_m \neq 0, \mathrm{A}_{\mathrm{S}',i} \neq 0\}.$$

On dira que M est bien gradué s'il satisfait la condition suivante : (BGr) : Si n et n' sont deux entiers tels que $n<n'$, $\mathrm{U}_n \neq 0$, $\mathrm{U}_{n'} \neq 0$, alors :

$$n' \in \mathrm{I}_n \Rightarrow \mathrm{I}_n \subset \mathrm{I}_{n'}$$
$$n' \notin \mathrm{I}_n \Rightarrow \mathrm{I}_n \cap \mathrm{I}_{n'} = \emptyset.$$

Soient $\mathrm{I}(1), \ldots, \mathrm{I}(h)$ les éléments maximaux (pour l'inclusion) de la famille $\{\mathrm{I}_n \mid n \in \mathbb{Z}\}$. Avec les notations précédentes on a :

5.4.2. Proposition. — *Supposons que* M *soit bien gradué.*

(i) *On a une décomposition canonique de* U *en somme directe :*

$$\mathrm{U} = \bigoplus_{k=1}^{h} \mathrm{U}^{(k)} \qquad \textit{avec} \quad \mathrm{U}^{(k)} = \bigoplus_{l \in \mathrm{I}_k} \mathrm{U}_l, \qquad 1 \leqq k \leqq h.$$

(ii) *Il existe une décomposition unique de* M *en somme directe de sous-modules gradués* $\mathrm{M} = \bigoplus_{k=1}^{h} \mathrm{M}^{(k)}$, *qui induit, modulo* $\mathscr{M}_{\mathrm{S}'}$, *la décomposition de* U *précédente.*

Démonstration. — Puisque M satisfait la condition (BGr), la famille $\{\mathrm{I}(k) \mid 1 \leqq k \leqq h\}$ est une partition de l'ensemble $\{n \mid \mathrm{U}_n \neq 0\}$, d'où (i).

Par ailleurs, si $\mathrm{M} = \bigoplus_{k} \mathrm{M}^{(k)}$ est une décomposition de M en somme directe telle que $\mathrm{M}^{(k)} / \mathscr{M}_{\mathrm{S}'} \mathrm{M}^{(k)} = \mathrm{U}^{(k)}$, alors $\mathrm{M}^{(k)}$ est le $\mathrm{A}_{\mathrm{S}'}$-sous-module de M engendré par les sous-groupes M_l avec $l \in \mathrm{I}(k)$; et il est clair que la famille des sous-modules $\mathrm{M}^{(k)}$ ainsi définie donne une décomposition de M en somme directe, d'où (ii).

5.4.3. Par inspection de la table 5.3, on vérifie que le module $\mathrm{M}_{\rho'}$ est bien gradué pour toute représentation irréductible ρ' de G′, à la seule exception de ρ'_3 du type $\mathbf{E}_7$ et ρ'_5 de $\mathbf{E}_8$. Or le type $\mathbf{E}_7$ est traité facilement à partir du type $\mathbf{E}_6$ (comme on verra dans le paragraphe 5.5) et le cas ρ_5 de $\mathbf{E}_8$ sera traité à part; donc, sauf indication contraire, on exclue dans la suite les cas exceptionnels. En conséquence de la proposition précédente, il y a une décomposition canonique de $\mathrm{M}_{\rho'}$ en somme directe de sous-modules $\mathrm{M}_{\rho'}^{(k)}$, avec $1 \leqq k \leqq h = h(\rho')$.

Supposons maintenant que ρ' apparaisse dans la décomposition en représentations irréductibles de $\operatorname{Ind} \rho$ où ρ est une représentation irréductible de G. Étant donné que l'opérateur θ_ρ est défini par la multiplication par le semi-invariant X de G′, on a l'inclusion : $\theta_\rho(\mathrm{M}_{\rho'}) \subset \mathrm{M}_{\varepsilon\rho'}$.

Notons $\theta_{\rho'} : \mathrm{M}_{\rho'} \to \mathrm{M}_{\varepsilon\rho'}$ la restriction de θ_ρ à $\mathrm{M}_{\rho'}$. On appelle *matrice* $[\theta_{\rho'}]$ associée à l'opérateur $\theta_{\rho'}$ le tableau indexé par $\{(k, \bar{k}), 1 \leqq k \leqq h(\rho'), 1 \leqq \bar{k} \leqq h(\varepsilon\rho')\}$ formé des blocks $[\theta_{\rho'}^{(k, \bar{k})}] : \mathrm{M}_{\rho'}^{(k)} \to \mathrm{M}_{\varepsilon\rho'}^{(\bar{k})}$.

5.4.4. *Remarques.* — (1) Notons $\mathrm{I}_{\rho'}$ l'ensemble de degrés $\{l \mid \mathrm{U}_{\rho', l} \neq 0\}$, et $\{\mathrm{I}_{\rho'}(k) \mid 1 \leqq k \leqq h\}$ la partition canonique de $\mathrm{I}_{\rho'}$ définie en 5.4.1. La table 5.3 montre que $\dim_{\mathbb{C}} \mathrm{U}_{\rho'}^{(k)}$ est égal à 1 ou 2 (et par suite $\operatorname{rang}_{\mathrm{A}_{\mathrm{S}'}} \mathrm{M}_{\rho'}^{(k)}$ est égal à 1 ou 2).

De fait on trouve trois cas possibles :

(*a*) $\mathrm{I}_{\rho'}(k) = \{l\}$ et $\dim_{\mathbb{C}} \mathrm{U}_{\rho', l} = 1$;

(*b*) $\mathrm{I}_{\rho'}(k) = \{l\}$ et $\dim_{\mathbb{C}} \mathrm{U}_{\rho', l} = 2$;

(*c*) $\mathrm{I}_{\rho'}(k) = \{l_1, l_2\}$ avec $l_1 < l_2$
et $\dim_{\mathbb{C}} \mathrm{U}_{\rho', l_1} = \dim_{\mathbb{C}} \mathrm{U}_{\rho', l_2} = 1$.

Dans le dernier cas on a une filtration de $\mathrm{M}_{\rho'}(k)$, $\mathrm{N}_{\rho'}^{(k)} \subset \mathrm{M}_{\rho'}^{(k)}$, où $\mathrm{N}_{\rho'}^{(k)}$ est le $\mathrm{A}_{\mathrm{S}'}$-sous-module de $\mathrm{M}_{\rho'}^{(k)}$ engendré par le sous-groupe M_{ρ', l_1}.

(2) En considérant les différents cas possibles [(*a*), (*b*) ou (*c*) de la remarque précédente] pour k et $\bar{k}$, on constate le fait suivant : Soit $B_{\rho'}^{(k)}$ une base, formée d'éléments homogènes, de $M_{\rho'}^{(k)}$. Alors il existe une base $B_{\varepsilon\rho'}^{(\bar{k})}$, formée d'éléments homogènes, de $M_{\varepsilon\rho'}^{(\bar{k})}$ telle que la matrice de $[\theta_{\rho'}^{(k,\bar{k})}]$ soit du type indiqué suivant le cas :

$\operatorname{rang}_{A_{S'}} M_{\rho'}^{(k)}$	$\operatorname{rang}_{A_{S'}} M_{\varepsilon\rho'}^{(\bar{k})}$	matrice de $[\theta_{\rho'}^{(k,\bar{k})}]$
1	1	$[f]$
1	2	$\begin{bmatrix} f \\ 0 \end{bmatrix}$
2	1	$[f \ \ 0]$
2	2	$\begin{bmatrix} f & h \\ 0 & g \end{bmatrix}$

où f (resp. g, resp. h) est un élément de $A_{S'}$ déterminé à un facteur scalaire près, ou nul.

(3) Posons $I_{S'} = \{ i \mid A_{S',i} \neq 0 \}$. Soit $l \in I_{\rho'}(k)$, et $\bar{l} \in I_{\rho'}(\bar{k})$. Considérons un élément m (resp. $\bar{m}$) homogène de degré l (resp. $\bar{l}$) appartenant à une base homogène $B_{\rho'}^{(k)}$ (resp. $B_{\varepsilon\rho'}^{(\bar{k})}$) de $M_{\rho'}^{(k)}$ (resp. de $M_{\varepsilon\rho'}^{(\bar{k})}$).

Alors, la composante de $\theta_{\rho'}(m)$ suivant $\bar{m}$ est un polynôme de la forme $\sum_{(a,b)} \alpha_{a,b} Y^a Z^b$, où a, b est un couple d'entiers positifs ou nuls parcourant l'ensemble :

$$\{ (a, b) \mid a \deg Y + b \deg Z = l - \bar{l} + \deg X \},$$

et où les $\alpha_{a,b}$ sont des constantes.

On en déduit donc, que si $(l - \bar{l} + \deg X) \notin I_{S'}$, alors ce polynôme est nul. De plus, sauf pour quelques représentations dans les cas $\mathbf{D}_n$ et $\mathbf{E}_7$, les polynômes en question sont des monômes.

(4) Soit ρ une représentation autoduale de G, et $\rho' \oplus \varepsilon\rho' = \operatorname{Ind} \rho$ la décomposition de $\operatorname{Ind} \rho$ en représentations irréductibles de G′ (prop. 3.4). Alors la décomposition $M_\rho = M_{\rho'} \oplus M_{\varepsilon\rho'}$ (cor. 3.6) induit une décomposition antidiagonale de θ_ρ, i. e. de la forme :

$$\begin{bmatrix} 0 & [\theta_{\varepsilon\rho'}] \\ [\theta_{\rho'}] & 0 \end{bmatrix}$$

(5) Soient ρ une représentation irréductible de G et $(x_i)_{i\in I}$ une famille finie d'éléments homogènes de A_V tels que pour tout i, $\mathbb{C}[G]x_i$ soit un sous-G-module simple de A_V de type ρ et tels que pour tout couple (i, j) il existe un $\mathbb{C}[G]$-isomorphisme $u_{i,j} : \mathbb{C}[G]x_i \to \mathbb{C}[G]x_j$ qui envoie x_i sur x_j. Alors il existe une famille $\langle x_i \rangle$, $i \in I$, d'éléments homogènes de $M_\rho = \operatorname{Hom}_{\mathbb{C}[G]}(E_\rho, A_V)$ tels que :
(*a*) $\langle x_i \rangle (E_\rho) = \mathbb{C}[G]x_i$.

(*b*) Le diagramme suivant soit commutatif :

$$\begin{tikzcd}
 & \mathbb{C}[G]\, x_i \arrow[dd, "u_{i.j}"] \\
E_\rho \arrow[ur, "\langle x_i\rangle"] \arrow[dr, "\langle x_j\rangle"'] & \\
 & \mathbb{C}[G]\, x_j
\end{tikzcd}$$

Une telle famille est déterminée au choix d'un scalaire non nul près. La même remarque est valable si on remplace G par G′.

(6) Étant donné que X^2 est un invariant du groupe G′, on a une relation de la forme :

$$(\star) \qquad X^2 = f(Y, Z)$$

où f est un polynôme à coefficients constants. On retrouve ainsi le fait que les points singuliers considérés sont des points doubles. Soient $\langle 1 \rangle$, $\langle X \rangle$ les éléments de la base de M_{ρ_0} (comme $A_{S'}$-module) déterminés par le polynôme constant 1 et par X respectivement. Avec cette base, la matrice de θ_{ρ_0} est la suivante :

$$\begin{bmatrix} 0 & f(Y, Z) \\ 1 & 0 \end{bmatrix}$$

(7) Comme conséquence de la proposition 4.4 (i), on a l'égalité :

$$(\theta_\rho)^2 = f(Y, Z) . \mathrm{Id}_{M_\rho},$$

où $f(Y, Z)$ provient de l'équation $(\star)$ de la remarque précédente. En particulier, si ρ est autoduale [*voir* remarque (4)], on a :

$$[\theta_{\rho'}]\,[\theta_{\varepsilon\rho'}] = f(Y, Z)\, \mathrm{Id}_{M_{\varepsilon\rho'}},$$
$$[\theta_{\varepsilon\rho'}]\,[\theta_{\rho'}] = f(Y, Z)\, \mathrm{Id}_{M_{\rho'}}$$

(8) Soient ρ_i, ρ_j et ρ_k trois représentations de G telles qu'on ait $\rho_i \otimes \rho_j = \rho_k$. Alors on a l'inclusion $M_{\rho_i} . M_{\rho_j} \subset M_{\rho_k}$, où $M_{\rho_i} . M_{\rho_j}$ dénote le produit symétrique de M_{ρ_i} avec M_{ρ_j}. Par suite, si m_i (resp. m_j) est un élément de M_{ρ_i} (resp. de M_{ρ_j}) on a les égalités :

$$\theta_{\rho_k}(m_i . m_j) = (\theta_{\rho_i}(m_i)) . m_j = m_i . (\theta_{\rho_j}(m_j)).$$

5.5. CALCUL DE LA RÉSOLUTION G-ÉQUIVARIANTE DE A_V (PROP. 4.4). — On utilisera la notation du paragraphe 3.7 pour les représentations. Les matrices données dans la liste suivante sont associées à des bases dont les éléments sont ordonnés par degrés croissants (*voir* table 5.3), pour chaque module $M_{\rho'}$. Les calculs sont effectués en utilisant les remarques 5.4.4. Pour chaque matrice on donne ses coordonnées en précisant la valeur du facteur constant seulement dans les cas où cette information est utile pour la suite, i. e. si la coordonnée en question n'est pas un monôme, ou si la représentation n'est pas autoduale; dans les autres cas on indique par $\star$ un coefficient constant non nul.

5.5.1. *Type* $\mathbf{A}_n$ $(n \geqq 1)$. — Dans ce cas il est plus facile de faire un calcul explicite.

Soit G le sous-groupe de SL(2, $\mathbb{C}$) engendré par $g=\begin{bmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{bmatrix}$, où $\xi = e^{2\,i\pi/(n+1)}$.

Dans l'algèbre $A_V = \mathbb{C}[x, y]$ les polynômes $Y = x^{n+1} + y^{n+1}$ et $Z = xy$ engendrent $A_{S'}$, et le jacobien de Y et Z est égal (à un facteur scalaire près) à $(x^{n+1} - y^{n+1})$.

Pour simplifier les calculs, on considère les invariants de A_V pour l'action de G, définis par :

$$\overline{X} = x^{n+1}, \qquad \overline{Y} = y^{n+1}, \qquad \overline{Z} = Z = xy.$$

On notera que $\overline{X}$ et $\overline{Y}$ ne sont pas des invariants sous G′ et qu'on a $X = \overline{X} - \overline{Y}$ à un facteur près.

Notons $\rho_j (1 \leqq j \leqq n)$ la représentation irréductible (à isomorphisme près) dont le caractère en g vaut :

$$\chi_{\rho_j}(g) = \xi^j.$$

Posons $s_1(j) = \langle x^j \rangle$, $s_2(j) = \langle y^{n+1-j} \rangle$; ils forment une base de M_{ρ_j} [*cf.* remarque (5) de 5.4.4], $1 \leqq j \leqq n$.

Alors on a les relations :

$$\text{(R)} \qquad \begin{cases} \overline{Y}\, s_1(j) = \overline{Z}^j s_2(j) \\ \overline{Z}^{(n+1-j)} s_1(j) = \overline{X}\, s_2(j) \end{cases} \qquad (1 \leqq j \leqq n).$$

La relation $\overline{X}\,\overline{Y} = \overline{Z}^{(n+1)}$, qui n'est autre chose que l'équation, en termes de $\overline{X}$, $\overline{Y}$, $\overline{Z}$, de la singularité quotient, détermine avec les relations (R) la résolution de A_V (comme $A_S[G]$-module) de la proposition 4.4.

5.5.2. *Type* $\mathbf{D}_n (n \geqq 4)$. — Dans ce cas aussi on peut faire les calculs explicitement, en considérant le groupe dihédral binaire G engendré par :

$$g = \begin{bmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{bmatrix} \quad \text{avec} \quad \xi = e^{2\,i\pi/2\,(n-2)} \quad \text{et par} \quad \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Soient $Y = x^{2\,(n-2)} + (-1)^n y^{2\,(n-2)}$, $Z = x^2 y^2$ deux générateurs de $A_{S'}$, et notons X le polynôme $xy(x^{2\,(n-2)} + (-1)^{n-1} y^{2\,(n-2)})$, qui est égal au jacobien de Y et Z à un facteur scalaire près. On détermine une base de M_ρ, $\rho \in \operatorname{Irr}(G)$, en utilisant le cas cyclique (5.5.1) et le fait que G contient le sous-groupe distingué cyclique d'ordre $2(n-2)$ engendré par g. Avec les notations du paragraphe 3.7 pour les représentations de G on trouve :

(a) ρ_1 :

$$[\theta_{\rho_1}] = \begin{bmatrix} 0 & Y^2 + 4(-1)^{(n-1)} Z^{(n-2)} \\ Z & 0 \end{bmatrix}.$$

[Avec les bases de $M_{\varepsilon\rho'_1}$ et $M_{\rho'_1}$ respectivement : $\langle xy \rangle$, $\langle x^{2n-4} + (-1)^{n-1} y^{2n-4} \rangle$].

(*b*) ρ_j, $2 \leqq j \leqq n-2$:

j pair :

$$[\theta_{\rho_j}] = \begin{bmatrix} 0 & 0 & \star YZ & \star Z^{(n-1-j/2)} \\ 0 & 0 & \star Z^{j/2} & \star Y \\ \star Y & \star Z^{n-1-j/2} & 0 & 0 \\ \star Z^{j-2} & \star YZ & 0 & 0 \end{bmatrix}$$

(Avec les bases :
de $M_{\rho'_j}$: $\langle x^{j-1} \rangle$, $\langle xy^{2n-j-2} \rangle$;
de $M_{\varepsilon\rho'_j}$: $\langle x^j y \rangle$, $\langle y^{2n-j-3} \rangle$).

j impair :

$$[\theta_{\rho_j}] = \begin{bmatrix} 0 & 0 & \star YZ & \star Z^{(n-1-(j-1)/2)} \\ 0 & 0 & \star Z^{(j+1)/2} & \star YZ \\ \star Y & \star Z^{(n-2-(j-1)/2)} & 0 & 0 \\ \star Z^{(j-1)/2} & \star Y & 0 & 0 \end{bmatrix}$$

(Avec les bases :
de $M_{\rho'_j}$: $\langle x^{j-1} \rangle$, $\langle y^{2n-j-3} \rangle$;
de $M_{\varepsilon\rho'_j}$: $\langle x^j y \rangle$, $\langle xy^{2n-j-2} \rangle$).

(*c*) ρ_{n-1} :

n pair :

$$[\theta_{\rho_{n-1}}] = \begin{bmatrix} 0 & Y + 2Z^{(n-2)/2} \\ YZ - 2Z^{n/2} & 0 \end{bmatrix}.$$

[Avec les bases :
de $M_{\rho'_{n-1}}$: $\langle xy(x^{n-2} + (-1)^{n-1} y^{n-2}) \rangle$;
de $M_{\varepsilon\rho'_{n-1}}$: $\langle x^{n-2} + (-1)^j y^{n-2} \rangle$].

n impair :

$$[\theta_{\rho_{n-1}}] = \begin{bmatrix} YZ & -2Z^{(n-1)/2} \\ 2Z^{(n-1)/2} & Y \end{bmatrix}.$$

(même base).

(*d*) ρ_n :

n pair :

$$[\theta_{\rho_n}] = \begin{bmatrix} 0 & Y - 2Z^{(n-2)/2} \\ YZ + 2Z^{n/2} & 0 \end{bmatrix}.$$

[Avec les bases :
de $M_{\rho'_n}$: $\langle xy(x^{n+2} + (-1)^n y^{n-2}) \rangle$;
de $M_{\varepsilon\rho'_n}$: $\langle x^{n-2} + (-1)^{n-1} y^{n-2} \rangle$].

n impair :

$$[\theta_{\rho_n}]=\begin{bmatrix} YZ & 2Z^{(n-1)/2} \\ -2Z^{(n-1)/2} & Y \end{bmatrix}$$

(même base).

5.5.3. *Type* $\mathbf{E}_6$

Quitte à multiplier X, Y, Z par des facteurs scalaires, on peut normaliser l'équation de S de sorte qu'on ait :

$$X^2=Y^3+Z^4.$$

Les représentations ρ_1 et ρ_2 sont autoduales. On peut choisir des bases ordonnées par ordre ascendant des degrés telles que les matrices soient :

$$[\theta_{\rho'_1}]=\begin{bmatrix} \star Y & \star Z^3 \\ \star Z & \star Y^2 \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_1}]=\begin{bmatrix} \star Y^2 & \star Z^3 \\ \star Z & \star Y \end{bmatrix},$$

$$[\theta_{\rho'_2}]=\begin{bmatrix} \star Y & 0 & \star Z^2 \\ \star Z & \star Y & 0 \\ 0 & \star Z & \star Y \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_2}]=\begin{bmatrix} \star Y^2 & \star Z^3 & \star YZ^2 \\ \star YZ & \star Y^2 & \star Z^3 \\ \star Z^2 & \star YZ & \star Y^2 \end{bmatrix},$$

Pour les représentations ρ_3, $\rho_3^\vee$, ρ_4, $\rho_4^\vee$ on trouve :

$$[\theta_{\rho_3}]=\begin{bmatrix} \alpha Z^2 & 0 & Y^2 & 0 \\ 0 & \alpha Z^2 & \beta YZ & Y^2 \\ Y & 0 & -\alpha Z^2 & 0 \\ -\beta Z & Y & 0 & -\alpha Z^2 \end{bmatrix},$$

$$[\theta_{\rho_3^\vee}]=\begin{bmatrix} -\alpha Z^2 & 0 & Y^2 & 0 \\ 0 & -\alpha Z^2 & \beta YZ & Y^2 \\ Y & 0 & \alpha Z^2 & 0 \\ -\beta Z & Y & 0 & \alpha Z^2 \end{bmatrix},$$

$$[\theta_{\rho_4}]=\begin{bmatrix} \alpha Z^2 & Y^2 \\ Y & -\alpha Z^2 \end{bmatrix}, \qquad [\theta_{\rho_4^\vee}]=\begin{bmatrix} -\alpha Z^2 & Y^2 \\ Y & \alpha Z^2 \end{bmatrix},$$

où α et β sont deux constantes non nulles, $\alpha^2=1$ (on peut déterminer la valeur de β, mais c'est inutile pour la suite).

En fait le choix $\alpha=1$ (ou $\alpha=-1$) précise la représentation de G notée ρ_3 (et par suite détermine les trois autres).

On a utilisé en particulier le fait qu'on a les égalités :

$$\rho_1\otimes\rho_4=\rho_3 \qquad \text{et} \qquad \rho_4\otimes\rho_4^\vee=\rho_0$$

et la remarque (8) de 5.4.4.

A titre d'exemple, on montre comment on détermine les matrices dans le cas ρ_1 :

Notons $\{(1),(11)\}$ [resp. $\{(5),(7)\}$] une base de $M_{\rho_1'}$ (resp. de $M_{\varepsilon\rho_1'}$) dont les éléments sont de degré égal au chiffre indiqué (*voir* table 5.3.3).

Par des considérations de degrés [remarque (3) de 5.4.4], en considérant les degrés de X, Y, Z (*voir* § 3.2), les matrices associées aux opérateurs $\theta_{\rho_1'} : M_{\rho_1'} \to M_{\varepsilon\rho_1'}$ et $\theta_{\varepsilon\rho_1'} \to M_{\rho_1'}$ avec ces bases, sont de la forme :

$$[\theta_{\rho_1'}]=\begin{array}{c} \begin{array}{cc} \theta(1) & \theta(11) \end{array} \\ \begin{bmatrix} aY & bZ^3 \\ cZ & dY^2 \end{bmatrix} \end{array} \begin{array}{l} \\ (5) \\ (7), \end{array} \qquad [\theta_{\varepsilon\rho_1'}]=\begin{array}{c} \begin{array}{cc} \theta(5) & \theta(7) \end{array} \\ \begin{bmatrix} eY^2 & fZ^3 \\ gZ & hY \end{bmatrix} \end{array} \begin{array}{l} \\ (1) \\ (11) \end{array}$$

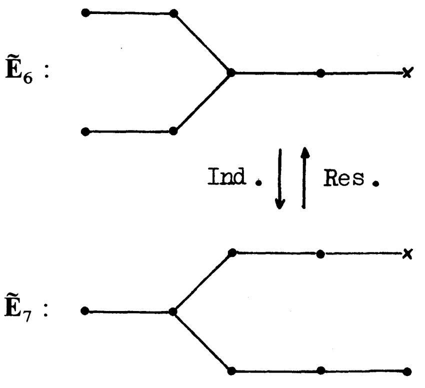
où $a, b, \ldots, h$ sont des coefficients constants. Or ces matrices doivent satisfaire les égalités :

$$[\theta_{\rho_1'}][\theta_{\varepsilon\rho_1'}]=(Y^3+Z^4)\,\mathrm{Id}_{M_{\varepsilon\rho_1'}} \quad \text{et} \quad [\theta_{\varepsilon\rho_1'}][\theta_{\rho_1'}]=(Y^3+Z^4)\,\mathrm{Id}_{M_{\rho_1'}}$$

[remarque (7) de 5.4.4], d'où en particulier : $ae=bg=ef=dh=1$, et par suite tous les coefficients sont non nuls. En fait on peut être plus précis, car en faisant des homothéties sur les éléments des bases, on peut supposer $a=b=c=1$, d'où $e=f=g=1$. D'autre part on a aussi $1+h=af+bh=0$, d'où $h=-1$ et par suite $d=-1$:

$$\begin{bmatrix} Y & Z^3 \\ Z & -Y^2 \end{bmatrix}, \quad \begin{bmatrix} Y^2 & Z^3 \\ Z & -Y \end{bmatrix}.$$

5.5.4. *Type* $\mathbf{E}_7$. — Un groupe G de type $\mathbf{E}_7$ contient un sous-groupe distingué d'indice 2 de type $\mathbf{E}_6$ (ceci est dû essentiellement au fait que la réunion des sommets d'un tétraèdre et du tétraèdre dual est l'ensemble de sommets d'un cube). La relation (par induction ou restriction) entre les représentations irréductibles de l'un et de l'autre est exprimée schématiquement par la correspondance de diagrammes suivante :



de sorte que la restriction (resp. l'induction) d'une représentation irréductible du groupe de type $\tilde{\mathbf{E}}_7$ (resp. de type $\tilde{\mathbf{E}}_6$) correspondante à un sommet du diagramme $\tilde{\mathbf{E}}_7$ (resp. de $\tilde{\mathbf{E}}_6$) est la somme de représentations correspondantes aux sommets du diagramme $\tilde{\mathbf{E}}_6$ (resp. $\tilde{\mathbf{E}}_7$) qui se trouvent sur la même verticale.

D'autre part, si on note Y_6, Z_6 (resp. Y_7, Z_7) des générateurs de $A_{S'}$ et X_6 (resp. X_7) leur jacobien dans le cas $\mathbf{E}_6$ (resp. $\mathbf{E}_7$), on a les égalités (à un facteur constant non nul près) suivantes :

$$X_7 = X_6 Z_6, \qquad Y_7 = Z_6^2, \qquad Z_7 = Y_6.$$

Ces deux remarques permettent de déduire la résolution équivariante de A_V pour le type $\mathbf{E}_7$ à partir de celle de type $\mathbf{E}_6$.

Quitte à multiplier $X = X_7$, $Y = Y_7$ et $Z = Z_7$ par des constantes, on a la relation :

$$X^2 = Y(Y^2 + Z^3).$$

Les matrices trouvées sont les suivantes :

$$[\theta_{\rho'_1}] = \begin{bmatrix} \star Y & \star YZ^2 \\ \star Z & \star Y^2 \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_1}] = \begin{bmatrix} \star Y^2 & \star YZ^2 \\ \star Z & \star Y \end{bmatrix},$$

$$[\theta_{\rho'_2}] = \begin{bmatrix} \star Y & 0 & \star YZ \\ \star Z & \star Y & 0 \\ 0 & \star Z & \star Y \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_2}] = \begin{bmatrix} \star Y^2 & \star YZ^2 & \star Y^2 Z \\ \star YZ & \star Y^2 & \star YZ^2 \\ \star Z^2 & \star YZ & \star Y^2 \end{bmatrix},$$

$$[\theta_{\rho'_3}] = \begin{bmatrix} 0 & \star YZ & \star Y^2 & \star YZ^2 \\ \star Y & 0 & \star YZ & \star Y^2 \\ \star Y & \star Z^2 & 0 & 0 \\ \star Z & \star Y & 0 & 0 \end{bmatrix},$$

$$[\theta_{\varepsilon\rho'_3}] = \begin{bmatrix} 0 & 0 & \star Y^2 & \star YZ^2 \\ 0 & 0 & \star YZ & \star Y^2 \\ \star Y & \star Z^2 & 0 & \star YZ \\ \star Z & \star Y & \star Y & 0 \end{bmatrix}$$

$$[\theta_{\rho'_4}] = \begin{bmatrix} 0 & \star YZ & \star Y^2 \\ \star Y & 0 & \star YZ \\ \star Z & \star Y & 0 \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_4}] = \begin{bmatrix} \star YZ & \star Y^2 & \star YZ^2 \\ \star Z^2 & \star YZ & \star Y^2 \\ \star Y & \star Z^2 & \star YZ \end{bmatrix},$$

$$[\theta_{\rho'_5}] = \begin{bmatrix} \star Z^2 & \star Y \\ \star Y & \star YZ \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_5}] = \begin{bmatrix} \star YZ & \star Y^2 \\ \star Y & \star Z^2 \end{bmatrix},$$

$$[\theta_{\rho'_6}] = [Y], \qquad [\theta_{\varepsilon\rho'_6}] = [Y^2 + Z^3],$$

$$[\theta_{\rho'_7}] = \begin{bmatrix} \star Y & \star Z^2 \\ \star Z & \star Y \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_7}] = \begin{bmatrix} \star Y^2 & \star YZ^2 \\ \star YZ & \star Y^2 \end{bmatrix},$$

5.5.5. *Type* $\mathbf{E}_8$. — Quitte à multiplier X, Y, Z par des facteurs constants, on a la relation :

$$X^2 = Y^3 + Z^5$$

Les matrices trouvées sont les suivantes :

$$[\theta_{\rho'_1}] = \begin{bmatrix} \star Y & \star Z^4 \\ \star Z & \star Y^2 \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_1}] = \begin{bmatrix} \star Y^2 & \star Z^4 \\ \star Z & \star Y \end{bmatrix},$$

$$[\theta_{\rho'_2}] = \begin{bmatrix} \star Y & 0 & \star Z^3 \\ \star Z & \star Y & 0 \\ 0 & \star Z & \star Y \end{bmatrix}, \qquad [\theta_{\varepsilon\rho'_2}] = \begin{bmatrix} \star Y^2 & \star Z^4 & \star YZ^3 \\ \star YZ & \star Y^2 & \star Z^4 \\ \star Z^2 & \star YZ & \star Y^2 \end{bmatrix},$$

$$[\theta_{\rho'_3}] = \begin{bmatrix} \star Z^2 & 0 & \star Y^2 & \star Z^4 \\ \star Y & 0 & \star Z^3 & \star YZ^2 \\ 0 & \star Z^2 & \star YZ & \star Y^2 \\ \star Z & \star Y & 0 & \star Z^3 \end{bmatrix},$$

$$[\theta_{\varepsilon\rho'_3}] = \begin{bmatrix} \star Z^3 & \star Y^2 & \star YZ^2 & \star Z^4 \\ 0 & \star YZ & \star Z^3 & \star Y^2 \\ \star Y & \star Z^2 & 0 & 0 \\ \star Z & 0 & \star Y & \star Z^2 \end{bmatrix},$$

$$[\theta_{\rho'_4}] = \begin{bmatrix} \star Z^2 & 0 & 0 & 0 & \star Y^2 \\ \star Y & 0 & 0 & 0 & \star Z^3 \\ 0 & \star Y & \star Z^2 & 0 & 0 \\ \star Z & 0 & \star Y & \star Z^2 & 0 \\ 0 & \star Z & 0 & \star Y & \star Z^2 \end{bmatrix},$$

$$[\theta_{\varepsilon\rho'_4}] = \begin{bmatrix} \star Z^3 & \star Y^2 & 0 & 0 & 0 \\ 0 & \star Z^3 & \star Y^2 & \star YZ^2 & \star Z^4 \\ 0 & \star YZ & \star Z^3 & \star Y^2 & \star YZ^2 \\ \star Z^2 & 0 & \star YZ & \star Z^3 & \star Y^2 \\ \star Y & \star Z^2 & 0 & 0 & 0 \end{bmatrix},$$

$$[\theta_{\rho'_5}]=\begin{bmatrix} 0 & 0 & \bigstar Z^3 & \bigstar Y^2 & \bigstar YZ^2 & \bigstar Z^4 \\ \bigstar Z^2 & 0 & 0 & \bigstar Z^3 & \bigstar Y^2 & \bigstar YZ^2 \\ 0 & \bigstar Z^2 & 0 & \bigstar YZ & \bigstar Z^3 & \bigstar Y^2 \\ \bigstar Y & \bigstar Z^2 & 0 & 0 & 0 & 0 \\ 0 & \bigstar Y & \bigstar Z^2 & 0 & 0 & 0 \\ \bigstar Z & 0 & \bigstar Y & 0 & 0 & 0 \end{bmatrix},$$

$$[\theta_{\varepsilon\rho'_5}]=\begin{bmatrix} 0 & 0 & 0 & \bigstar Y^2 & \bigstar YZ^2 & \bigstar Z^4 \\ 0 & 0 & 0 & \bigstar Z^3 & \bigstar Y^2 & \bigstar YZ^2 \\ 0 & 0 & 0 & \bigstar YZ & \bigstar Z^3 & \bigstar Y^2 \\ \bigstar Y & \bigstar Z^2 & 0 & 0 & 0 & \bigstar Z^3 \\ 0 & \bigstar Y & \bigstar Z^2 & \bigstar Z^2 & 0 & 0 \\ \bigstar Z & 0 & \bigstar Y & 0 & \bigstar Z^2 & 0 \end{bmatrix},$$

$$[\theta_{\rho'_6}]=\begin{bmatrix} 0 & \bigstar Z^3 & \bigstar Y^2 & \bigstar YZ \\ \bigstar Z^2 & 0 & 0 & \bigstar Y^2 \\ \bigstar Y & 0 & 0 & \bigstar Z^3 \\ \bigstar Z & \bigstar Y & \bigstar Z & 0 \end{bmatrix},$$

$$[\theta_{\varepsilon\rho'_6}]=\begin{bmatrix} O & \bigstar Z^3 & \bigstar Y^2 & 0 \\ \bigstar Z^2 & 0 & \bigstar YZ & \bigstar Y^2 \\ \bigstar Y & \bigstar Z^2 & 0 & \bigstar Z^3 \\ 0 & \bigstar Y & \bigstar Z^2 & 0 \end{bmatrix},$$

$$[\theta_{\rho'_7}]=\begin{bmatrix} \bigstar Z^2 & \bigstar Y^2 \\ \bigstar Y & \bigstar Z^3 \end{bmatrix},\qquad [\theta_{\varepsilon\rho'_7}]=\begin{bmatrix} \bigstar Z^3 & \bigstar Y^2 \\ \bigstar Y & \bigstar Z^2 \end{bmatrix},$$

$$[\theta_{\rho'_8}]=\begin{bmatrix} \bigstar Y & \bigstar Z^2 & 0 \\ 0 & \bigstar Y & \bigstar Z^2 \\ \bigstar Z & \bigstar Z^2 & \bigstar Y^2 \end{bmatrix},\qquad [\theta_{\varepsilon\rho'_8}]=\begin{bmatrix} \bigstar Y^2 & \bigstar YZ^2 & \bigstar Z^4 \\ \bigstar Z^3 & \bigstar Y^2 & \bigstar YZ^2 \\ \bigstar YZ & \bigstar Z^3 & \bigstar Y^2 \end{bmatrix},$$

Les matrices correspondantes aux indices 1, 2, 7 et 8 sont déterminées par des considérations de degrés [remarque (3), § 5.4.4] et par la condition de la remarque (7), § 5.4.4. La même méthode détermine les autres matrices, à quelques coefficients près; et ces derniers sont déterminés par la connaissance des matrices d'indices 1, 2, 7, 8 et la remarque (8) appliquée (par exemple) aux égalités :

$$\rho_1\otimes\rho_8=\rho_5,\qquad \rho_1\otimes\rho_7=\rho_6,\qquad \rho_1\otimes\rho_2=\rho_1\oplus\rho_3,\qquad \rho_1\otimes\rho_3=\rho_2\oplus\rho_4.$$

6. Désingularisations minimales; démonstration du théorème 2.2

6.1. On dispose de désingularisations minimales explicites $q : \tilde{S} \rightarrow S$ qui sont construites par une procédure de « soudure » à partir de plusieurs copies de $\mathbb{C}^2$ ([G], *voir* aussi [H], [L]). Le morphisme q qui contracte le diviseur exceptionnel $D = \bigcup_k d_k$ de $\tilde{S}$ est déterminé essentiellement par la donnée des cycles (à support dans D) de la forme $\sum_n n_k d_k$, qui seront représentés par des diagrammes analogues au graphe dual associé à D, où on remplace le sommet correspondant à une composante irréductible d_k par la multiplicité n_k.

6.1.1. *Type* $\mathbf{A}_n$ $(n \geqq 1)$. — Pour chaque entier k, $0 \leqq k \leqq n$, soit $\mathcal{U}_k$ une copie de $\mathbb{C}^2$ de coordonnées (u_k, v_k). Soit $\tilde{S}$ la surface (lisse) obtenue en identifiant les ouverts $\mathcal{U}_{k-1} - \{u_{k-1} = 0\}$ et $\mathcal{U}_k - \{v_k = 0\}$, $1 \leqq k \leqq n$, par les recollements :

$$u_{k-1} = 1/v_k, \qquad v_{k-1} = u_k v_k^2.$$

Soit d_k la courbe rationnelle lisse définie par $v_{k-1} = 0$ dans $\mathcal{U}_{k-1}$ et par $u_k = 0$ dans $\mathcal{U}_k$, $1 \leqq k \leqq n$. Alors chaque d_k a une auto-intersection -2 et le graphe dual associé à $D = \bigcup_k d_k$ est $\mathbf{A}_n$. Notons $C(\overline{X})$, $C(\overline{Y})$, $C(\overline{Z})$ les cycles à support dans D définis par les diagrammes :

$$C(\overline{X}) : n(n-1)\ldots 1 \left[\text{i.e. } C(\overline{X}) = \sum_{k=1}^{n} (n+1-k)\, d_k\right]$$

$$C(\overline{Y}) : 1\ \ 2\ldots n \left[\text{i.e. } C(\overline{Y}) = \sum_{k=1}^{n} k\, d_k\right]$$

$$C(\overline{Z}) : 1\ \ 1\ldots 1 \left[\text{i.e. } C(\overline{Z}) = \sum_{k=1}^{n} d_k\right].$$

A l'aide de ces cycles on définit une application $q : \tilde{S} \rightarrow \mathbb{C}^3$ par :

$$q|_{\mathcal{U}_k}(u_k, v_k) = (u_k^{(n+1-k)}\, v_k^{(n-k)},\ u_k^k\, v_k{}^{+1},\ u_k v_k), \qquad 0 \leqq k \leqq n.$$

Notons $(\overline{X}, \overline{Y}, \overline{Z})$ les coordonnées de $\mathbb{C}^3$. On vérifie que l'application q est une désingularisation minimale de la surface S (plongée dans $\mathbb{C}^3$) d'équation : $\overline{X}\,\overline{Y} = \overline{Z}^{(n+1)}$.

6.1.2. *Type* $\mathbf{D}_n$ $(n \geqq 4)$. — Pour chaque entier k, $0 \leqq k \leqq n$, soit $\mathcal{U}_k$ une copie de $\mathbb{C}^2$ de coordonnées (u_k, v_k). Soit $\tilde{S}$ la surface (lisse) obtenue en identifiant :

les ouverts $\mathcal{U}_{k-1} - \{u_{k-1} = 0\}$ et $\mathcal{U}_k - \{v_k = 0\}$, $1 \leqq k \leqq n-2$ par les recollements :

$$u_{k-1} = 1/v_k \quad \text{et} \quad v_{k-1} = u_k v_k^2,$$

les ouverts $\mathcal{U}_{n-2} - \{u_{n-2} = 0\}$ et $\mathcal{U}_{n-1} - \{v_{n-1} = 0\}$ par les recollements :

$$u_{n-2} = 1/v_{n-1} \quad \text{et} \quad \begin{cases} v_{n-2} + i = u_{n-1} v_{n-1}^2 & \text{si } n \text{ pair,} \\ v_{n-2} + 1 = u_{n-1} v_{n-1}^2 & \text{si } n \text{ impair,} \end{cases}$$

les ouverts $\mathscr{U}_{n-2}-\{u_{n-2}=0\}$ et $\mathscr{U}_{n}-\{v_{n}=0\}$ par les recollements :

$$u_{n-2}=1/v_n \qquad \text{et} \qquad \begin{cases} v_{n-2}-i=u_n v_n^2 & \text{si } n \text{ pair,} \\ v_{n-2}-1=u_n v_n^2 & \text{si } n \text{ impair,} \end{cases}$$

Soit d_k la courbe rationnelle lisse définie par $v_{k-1}=u_k=0$, avec $1 \leqq k \leqq n-2$;

d_{n-1} définie par :

$$u_{n-1}=0 \qquad \text{et} \qquad \begin{cases} v_{n-2}+i=0 & \text{si } n \text{ pair,} \\ v_{n-2}+1=0 & \text{si } n \text{ impair,} \end{cases}$$

d_n définie par :

$$u_n=0 \qquad \text{et} \qquad \begin{cases} v_{n-2}-i=0 & \text{si } n \text{ pair,} \\ v_{n-2}-1=0 & \text{si } n \text{ impair,} \end{cases}$$

Alors chaque d_k a une auto-intersection -2 et le graphe dual associé à $\mathrm{D}=\bigcup_k d_k$ est $\mathbf{D}_n$.

Avec les cycles C(X), C(Y), C(Z) définis par :

$$\begin{array}{rll} & n \text{ pair} & n \text{ impair} \\ \mathrm{C(X)}: & 2\ 3 \ldots (n-1)\begin{smallmatrix} n/2 \\ n/2 \end{smallmatrix} & 2\ 3 \ldots (n-1)\begin{smallmatrix} (n-1)/2 \\ (n-1)/2 \end{smallmatrix} \\ \mathrm{C(Y)}: & 1\ 2 \ldots (n-2)\begin{smallmatrix} n/2-1 \\ n/2-1 \end{smallmatrix} & 1\ 2 \ldots (n-2)\begin{smallmatrix} (n-1)/2 \\ (n-1)/2 \end{smallmatrix} \\ \mathrm{C(Z)}: & 2\ 2 \ldots 2_1^1 & 2\ 2 \ldots 2_1^1 \end{array}$$

On définit une application $q : \tilde{\mathrm{S}} \rightarrow \mathbb{C}^3$, dont les restrictions aux ouverts précédents sont données par :

si $0 \leqq k \leqq n-3$:

$$\mathrm{X}=u_k^{k+1} v_k^{k+2}(1+(-1)^n u_k^{2(n-k-2)} v_k^{2(n-k-3)})^{a_k},$$

$$\text{avec } a_k=\begin{cases} n/2 & \text{si } n \text{ pair,} \\ (n-1)/2 & \text{si } n \text{ impair,} \end{cases}$$

$$\mathrm{Y}=u_v^k v_k^{k+1}(1+(-1)^n u_k^{2(n-k-2)}) v_k^{2(n-k-3)})^{b_k},$$

$$\text{avec } b_k=\begin{cases} n/2-1 & \text{si } n \text{ pair,} \\ (n-1)/2 & \text{si } n \text{ impair,} \end{cases}$$

$$\mathrm{Z}=u_k^2 v_k^2(1+(-1)^n u_k^{2(n-k-2)} v_k^{2(n-k-3)}).$$

- -

$$\mathrm{X}=\begin{cases} u_{n-2}^{n-1}(1+v_{n-2}^2)^{n/2} & \text{(si } n \text{ pair),} \\ u_{n-2}^{n-1} v_{n-2}(v_{n-2}^2-1)^{(n-1)/2} & \text{(si } n \text{ impair),} \end{cases}$$

$$\mathrm{Y}=\begin{cases} u_{n-2}^{n-2} v_{n-2}(1+v_{n-2}^2)^{n/2-1} & \text{(si } n \text{ pair),} \\ u_{n-2}^{n-2}(v_{n-2}^2-1)^{(n-1)/2} & \text{(si } n \text{ impair),} \end{cases}$$

$$\mathrm{Z}=\begin{cases} u_{n-2}^2(1+v_{n-2}^2) & \text{(si } n \text{ pair),} \\ u_{n-2}^2(v_{n-2}^2-1) & \text{(si } n \text{ impair),} \end{cases}$$

$$X=\begin{cases}[u_{n-1}(u_{n-1}v_{n-1}^2-2i)]^{n/2}v_{n-1} & \text{si } n \text{ pair,}\\ [u_{n-1}(u_{n-1}v_{n-1}^2-2)]^{(n-1)/2}(u_{n-1}v_{n-1}^2-1) & \text{si } n \text{ impair,}\end{cases}$$

$$Y=\begin{cases}[u_{n-1}(u_{n-1}v_{n-1}^2-2i)]^{n/2-1}(u_{n-1}v_{n-1}^2-i & \text{si } n \text{ pair,}\\ [u_{n-1}(u_{n-1}v_{n-1}^2-2)]^{(n-1)/2}v_{n-1} & \text{si } n \text{ impair,}\end{cases}$$

$$Z=\begin{cases}u_{n-1}(u_{n-1}v_{n-1}^2-2i) & \text{si } n \text{ pair,}\\ u_{n-1}(u_{n-1}v_{n-1}^2-2) & \text{si } n \text{ impair,}\end{cases}$$

$$X=\begin{cases}[u_n(u_nv_n^2+2i)]^{n/2}v_n & \text{si } n \text{ pair,}\\ [u_n(u_nv_n^2+2)]^{(n-1)/2}(u_nv_n^2+1) & \text{si } n \text{ impair,}\end{cases}$$

$$Y=\begin{cases}[u_n(u_nv_n^2+2i)]^{n/2-1}(u_nv_n^2+i) & \text{si } n \text{ pair,}\\ [u_n(u_nv_n^2+2)]^{(n-1)/2} & \text{si } n \text{ impair,}\end{cases}$$

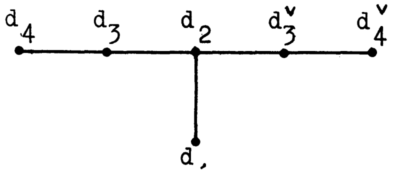
$$Z=\begin{cases}u_n(u_n v_n^2+2i) & \text{si } n \text{ pair,}\\ u_n(u_nv_n^2+2) & \text{si } n \text{ impair.}\end{cases}$$

L'application q est une désingularisation minimale de la surface S (plongée dans $\mathbb{C}^3$) d'équation $X^2=Z(Y^2+Z^{n-2})$.

6.1.3. *Type* $\mathbf{E}_6$. — Soient $\mathscr{U}_k$, $0\leqq k\leqq 4$, $\check{\mathscr{U}}_3$, $\check{\mathscr{U}}_4$ des copies de $\mathbb{C}^2$ de coordonnée (u_k, v_k), $0\leqq k\leqq 4$, $(\check{u}_3, \check{v}_3)$, $(\check{u}_4, \check{v}_4)$, respectivement. On considère la surface $\tilde{S}$ (lisse) (resp. la famille **des courbes rationnelles lisses** $\boldsymbol{d}$. **de** $\tilde{\mathbf{S}}$**) définie par les recollements (resp. par les équations) :**

$$\begin{array}{lll}
u_1=u_0^2v_0, & v_1=1/u_0, & d_1 : u_1=v_0=0,\\
u_2=u_0v_0^2, & v_2=1/v_0, & d_2 : u_2=u_0=0,\\
u_3=u_0^2(v_0+1), & v_3=1/u_0, & d_3 : u_3=v_0+1=0,\\
\check{u}_3=u_0^2(v_0-1), & \check{v}_3=1/u_0, & \check{d}_3 : \check{u}_3=v_0-1=0,\\
u_4=u_3^2v_3, & v_4=1/u_3, & d_4 : u_4=v_3=0,\\
\check{u}_4=\check{u}_3^2\check{v}_3, & \check{v}_4=1/\check{u}_3, & \check{d}_4 : \check{u}_4=\check{v}_3=0.
\end{array}$$

Chacune de ces courbes rationnelles a une auto-intersection -2, et le graphe dual associé à leur réunion est $\mathbf{E}_6$:



A l'aide des cycles définis par les diagrammes suivants :

$$C(X):\ \begin{matrix}2&4&6&4&2\\&&3&&\end{matrix},\qquad C(Y):\ \begin{matrix}2&3&4&3&2\\&&2&&\end{matrix},\qquad C(Z):\ \begin{matrix}1&2&3&2&1\\&&2&&\end{matrix}$$

on définit une application $q : \tilde{S}\rightarrow\mathbb{C}^3$ dont les restrictions aux ouverts précédents sont données par :

$$
\begin{aligned}
X &= u_0^6 v_0^3 (v_0^2-1)^4 = u_1^3 (u_1^2 v_1^4 - 1)^4 \\
&= u_2^6 v_2 (v_2^2-1)^4 = u_3^4 v_3^2 (u_3 v_3^2+1)^3 (u_3 v_3^2+2)^4 = u_4^2 (u_4^2 v_4^3+1)^3 (u_4^2 v_4^3+2)^4 \\
Y &= -u_0^4 v_0^2 (v_0^2-1)^3 = -u_1^2 (u_1^2 v_1^4-1)^3 \\
&= u_2^4 (v_2^2-1)^3 = -u_3^3 v_3^2 (u_3 v_3^2+1)^2 (u_3 v_3^2+2)^3 = -u_4^2 v_4 (u_4^2 v_4^3+1)^2 (u_4^2 v_4^3+2) \\
Z &= u_0^3 v_0^2 (v_0^2-1)^2 = u_1^2 v_1 (u_1^2 v_1^4-1)^2 \\
&= u_2^3 (v_2^2-1)^2 = u_3^2 v_3 (u_3 v_3^2+1)^2 (u_3 v_3^2+2)^2 = u_4 (u_4^2 v_4^3+1)^2 (u_4^2 v_4^3+2)^2.
\end{aligned}
$$

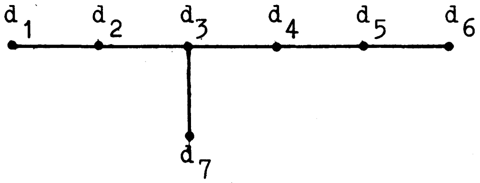
Pour l'ouvert $\check{\mathscr{U}}_3$ (resp. $\check{\mathscr{U}}_4$) on remplace u_3, v_3, $+1$, $+2$ (resp. u_4, v_4, $+1$, $+2$) dans l'expression correspondante à $\mathscr{U}_3$ (resp. à $\mathscr{U}_4$) par $\check{u}_3$, $\check{v}_3$, -1, -2 (resp. par $\check{u}_4$, $\check{v}_4$, -1, -2).

L'application q est une désingularisation minimale de la surface S (plongée dans $\mathbb{C}^3$) d'équation $X^2 = Y^3 + Z^4$.

6.1.4. *Type* $\mathbf{E}_7$. — Soient $\mathscr{U}_k$, $0 \leqq k \leqq 7$, des copies de $\mathbb{C}^2$ de coordonnées (u_k, v_k). On considère la surface $\tilde{S}$ (lisse) (resp. la famille des corbes rationnelles lisses d de $\tilde{S}$) définie par les recollements (resp. par les équations) :

$$
\begin{array}{llll}
u_k = v_{k-1} u_{k-1}^2, & v_k = 1/u_{k-1} & (1 \leqq k \leqq 6), & d_k : \ u_k = v_{k-1} = 0 \\
u_7 = (v_3+1) u_3^2, & v_7 = 1/u_3, & & d_7 : \ u_7 = v_3 + 1 = 0.
\end{array}
$$

Chacune de ces courbes rationnelles a une auto-intersection -2, et le graphe dual associé à leur réunion est $\mathbf{E}_7$:



A l'aide des cycles définis par les diagrammes suivants :

$$
C(X) : \begin{array}{cccccc} 3 & 6 & 9 & 7 & 5 & 3, \\ & & 5 & & & \end{array} \qquad
C(Y) : \begin{array}{cccccc} 2 & 4 & 6 & 5 & 4 & 3, \\ & & 3 & & & \end{array} \qquad
C(Z) : \begin{array}{cccccc} 2 & 3 & 4 & 3 & 2 & 1. \\ & & 2 & & & \end{array}
$$

On définit une application $q : \tilde{S} \to \mathbb{C}^3$ dont les restrictions aux ouverts précédents sont données par :

$$
\begin{aligned}
X &= v_0^3 (u_0^3 v_0^2+1)^5 = u_1^3 v_1^6 (u_1^2 v_1+1)^5 \\
&= u_2^6 v_2^9 (u_2+1)^5 = u_3^9 v_3^7 (1+v_3)^5 = u_4^7 v_4^5 (u_4 v_4^2+1)^5 \\
&= u_5^5 v_5^3 (u_5^2 v_5^3+1)^5 = u_6^3 v_6 (u_6^3 v_6^4+1)^5 = u_7^5 v_7 (u_7 v_7^2-1)^7,
\end{aligned}
$$

$$\begin{aligned}\mathrm{Y}&=v_0^2(u_0^3v_0^2+1)^3=u_1^2v_1^4(u_1^2v_1+1)^3\\&=u_2^4v_2^6(u_2+1)^3=u_3^6v_3^5(1+v_3)^3=u_4^5v_4^4(u_4v_4^2+1)^3\\&=u_5^4v_5^3(u_5^2v_5^3+1)^3=u_6^3v_6^2(u_6^3v_6^4+1)^3=u_7^3(u_7v_7^2-1)^5,\end{aligned}$$

$$\begin{aligned}\mathrm{Z}&=u_0v_0^2(u_0^3v_0^2+1)^2=u_1^2v_1^3(u_1^2v_1+1)^2\\&=u_2^3v_2^4(u_2+1)^2=u_3^4v_3^3(1+v_3)^2=u_4^3v_4^2(u_4v_4^2+1)^2\\&=u_5^2v_5(u_5^2v_5^3+1)^2=u_6(u_6^3v_6^4+1)^2=u_7^2(u_7v_7^2-1)^3.\end{aligned}$$

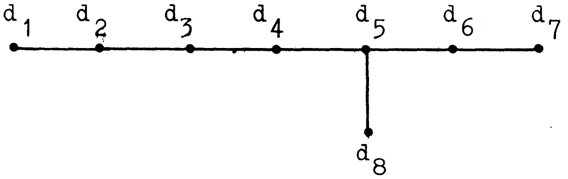
L'application q donne une résolution minimale de la surface S (plongée dans $\mathbb{C}^3$) d'équation $\mathrm{X}^2=\mathrm{Y}(\mathrm{Y}^2+\mathrm{Z}^3)$.

6.1.5. *Type* $\mathbf{E}_8$. — Soient $\mathcal{U}_k$, $0\leqq k\leqq 8$, des copies de $\mathbb{C}^2$ de coordonnées (u_k, v_k). On considère la surface $\tilde{\mathrm{S}}$ (lisse) (resp. la famille des courbes rationnelles lisses d de $\tilde{\mathrm{S}}$) définie par les recollements (resp. par les équations) :

$$u_k=v_{k-1}u_{k-1}^2,\qquad v_k=1/u_{k-1}\qquad(1\leqq k\leqq 7),\qquad d_k:\ u_k=v_{k-1}=0,$$

$$u_8=(v_5+1)u_5^2,\qquad v_8=1/u_5,\qquad d_8:\ u_8=v_5+1=0.$$

Chacune de ces courbes rationnelles a une auto-intersection -2, et le graphe dual associé à leur réunion est $\mathbf{E}_8$:



A l'aide des cycles définis par les diagrammes suivants :

$$\mathrm{C(X)}:\ \begin{matrix}3&6&9&12&15&10&5\\&&&&8&&\end{matrix},\qquad \mathrm{C(Y)}:\ \begin{matrix}2&4&6&8&10&7&4\\&&&&5&&\end{matrix}$$

$$\mathrm{C(Z)}:\ \begin{matrix}2&3&4&5&6&4&2\\&&&&3&&\end{matrix}$$

on définit une application $q:\tilde{\mathrm{S}}\to\mathbb{C}^3$ dont les restrictions aux ouverts précédents sont données par :

$$\begin{aligned}\mathrm{X}&=v_0^3(u_0^5v_0^4+1)^8=u_1^3v_1^6(u_1^4v_1^3+1)^8=u_2^6v_2^9(u_2^3v_2^2+1)^8=u_3^9v_3^{12}(u_3^2v_3+1)^8\\&=u_4^{12}v_4^{15}(u_4+1)^8=u_5^{15}v_5^{10}(v_5+1)^8=u_6^{10}v_6^5(u_6v_6^2+1)^8=u_7^5(u_7^2v_7^3+1)^8\\&=u_8^8v_8(u_8v_8^2-1)^{10},\end{aligned}$$

$$\begin{aligned}\mathrm{Y}&=v_0^2(u_0^5v_0^4+1)^5=u_1^2v_1^4(u_1^4v_1^3+1)^5=u_2^4v_2^6(u_2^3v_2^2+1)^5=u_3^6v_3^8(u_3^2v_3+1)^5\\&=u_4^8v_4^{10}(u_4+1)^5=u_5^{10}v_5^7(v_5+1)^5=u_6^7v_6^4(u_6v_6^2+1)^5=u_7^4v_7(u_7^2v_7^3+1)^5\\&=u_8^5(u_8v_8^2-1)^7,\end{aligned}$$

$$\begin{aligned}\mathrm{Z}=u_0 v_0^2(u_0^5 v_0^4+1)^3=u_1^2 v_1^3(u_1^4 v_1^3+1)^3=u_2^3 v_2^4(u_2^3 v_2^2+1)^3=u_3^4 v_3^5(u_3^2 v_3+1)^3\\ =u_4^5 v_4^6(u_4+1)^3=u_5^6 v_5^4(v_5+1)^3=u_6^4 v_6^2(u_6 v_6^2+1)^3=u_7^2(u_7^2 v_7^3+1)^3\\ =u_8^3(v_8 v_8-1)^4.\end{aligned}$$

L'application q est une désingularisation minimale de la surface S (plongée dans $\mathbb{C}^3$) d'équation $\mathrm{X}^2=\mathrm{Y}^3+\mathrm{Z}^5$.

6.2. DÉMONSTRATION DU THÉORÈME 2.2

6.2.1. *Démonstration de* (i) *: plan de la démonstration.* — Pour tout $\rho\in\mathrm{Irr}(\mathrm{G})$, considérons le fibré $\mathscr{F}_\rho$ (défini en paragraphe 2.7). On a $(rg\,\mathscr{F}_\rho, c_1(\mathscr{F}_\rho))=\pi(\rho)$ (prop. 1.2, cor. 2.9).

Nous allons montrer, cas par cas, qu'il existe un et un seul $d_\rho\in\mathrm{Irr}(\mathrm{D})$ tel que $\deg(\mathscr{F}_\rho|_{d_\rho})=1$ et $\deg(\mathscr{F}_\rho|_d)=0$ pour $d\neq d_\rho$, $d\in\mathrm{Irr}(\mathrm{D})$. Par inspection on vérifiera que $\rho\mapsto d_\rho$ est bijective et que $(d_{\rho_i}.\,d_{\rho_j})=a_{ij}$ pour $\rho_i\neq\rho_j$. On aura donc $c_1(\mathscr{F}_\rho)=b_{d_\rho}$ ce qui achèvera la démonstration de (i).

Le fibré $\mathscr{F}_\rho$ s'obtient en annulant la torsion de $q^*\mathscr{M}_\rho$, torsion qui est à support dans D. On a une résolution (induite par celle de 4.6) :

$$\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\mathscr{S}}\xrightarrow{\mathrm{T}_\rho^-}\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\mathscr{S}}\to\mathscr{M}_\rho\to 0$$

d'où une suite exacte sur $\tilde{\mathscr{S}}$:

$$\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\tilde{\mathscr{S}}}\xrightarrow{\tilde{\mathrm{T}}_\rho^-}\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\tilde{\mathscr{S}}}\to q^*\mathscr{M}_\rho\to 0$$

et par suite un homomorphisme surjectif :

$$\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\tilde{\mathscr{S}}}\xrightarrow{\overline{\varpi}_\rho}\mathscr{F}_\rho\to 0.$$

Les modules $\mathscr{M}_\rho$ sont des $\mathcal{O}_{\tilde{\mathscr{S}}}$-modules libres, et les homomorphismes $\mathrm{T}_\rho^-=\theta_\rho\otimes\mathrm{Id}_{\mathcal{O}_{\tilde{\mathscr{S}}}}-\mathrm{Id}_{\mathscr{M}_\rho}\otimes\mathrm{X}$ ont été décrits par des matrices (§ 5.5).

Sur chaque ouvert $\mathscr{U}_k$ de $\tilde{\mathscr{S}}$ on va décrire un sous-module libre L_k de $(\mathscr{M}_\rho\underset{\mathcal{O}_{\mathscr{S}'}}{\otimes}\mathcal{O}_{\tilde{\mathscr{S}}})|_{\mathscr{U}_k}$ tel que la projection $\varpi_\rho|_{\mathscr{U}_k}$ induise un isomorphisme de L_k sur L_k sur $\mathscr{F}_\rho|_{\mathscr{U}_k}$.

Le $c_1(\mathscr{F}_\rho)$ s'obtient alors en examinant les diviseurs sur D des matrices de changement de base de L_k à $\mathrm{L}_{k'}$, sur $\mathscr{U}_k\cap\mathscr{U}_{k'}$. Comme on connait explicitement les matrices décrivant T_ρ^- et l'application $q|_{\mathscr{U}_k}:\mathscr{U}_k\to\mathscr{S}$, le calcul est purement mécanique. On se bornera donc à traiter l'exemple de $\mathbf{A}_n$ et l'exemple de la représentation ρ_5 de $\mathbf{E}_8$ (correspondant au sommet central de $\mathbf{E}_8$).

6.2.2. *Démonstration de* (i), *type* $\mathbf{A}_n(n\geqq 1)$. — Soit j un entier, $1\leqq j\leqq n$. En vue de la définition de $q|_{\mathscr{U}_k}$ (*voir* § 6.1.1) on a $q^*\overline{\mathrm{Z}}^j|_{\mathscr{U}_k}=u_k^j v_k^j$, $q^*\overline{\mathrm{Y}}|_{\mathscr{U}_k}=u_k^k v_k^{k+1}$, $0\leqq k\leqq n$. Par conséquent, si $k \geqq j$, $q^* \overline{Z}^j |_{\mathscr{U}_k}$ divise $q^* \overline{Y} |_{\mathscr{U}_k}$, et si $k<j$, $q^* \overline{Y} |_{\mathscr{U}_k}$ divise $q^* \overline{Z}^j |_{\mathscr{U}_k}$; et par suite le $\mathscr{O}_{\mathscr{U}_k}$-module $\mathscr{F}_{\rho_j} |_{\mathscr{U}_k}$ est isomorphe au module libre de rang 1, L_k, engendré par :

$$\begin{array}{lll} s_1(j) & \text{si} & k \geqq j, \\ s_2(j) & \text{si} & k<j \end{array}$$

(*voir* § 5.5.1).

On a donc $\deg \mathscr{F}_{\rho_j} |_{d_k} = 0$ si $k \neq j$, $1 \leqq k \leqq n$, car les changements de base sur $\mathscr{U}_{k-1} \cap \mathscr{U}_k$ sont triviaux, et $\deg \mathscr{F}_{\rho_j} |_{d_j} = 1$, car le changement de base sur $\mathscr{U}_{j-1} \cap \mathscr{U}_j$ est : $s_2(j) = v_j s_1(j)$. Par suite $d_{\rho_j} = d_j$, $1 \leqq j \leqq n$, et $(d_i . d_j) = a_{ij}$, $1 \leqq i \neq j \leqq n$, d'où (i).

6.2.3. *Cas* ρ_5, *type* $\mathbf{E}_8$. — Avec la décomposition $M_{\rho_5} = M_{\rho'_5} \oplus M_{\varepsilon\rho'_5}$, l'homomorphisme $T^-_{\rho_5}$ est de la forme :

$$\begin{bmatrix} [-X \, \mathrm{Id}_{M_{\rho'_5}}] & [\theta_{\varepsilon\rho'_5}] \\ [\theta_{\rho'_5}] & [-X \, \mathrm{Id}_{M_{\varepsilon\rho'_5}}] \end{bmatrix}$$

où $[\theta_{\rho'_5}]$ et $[\theta_{\varepsilon\rho'_5}]$ ont été décrites en paragraphe 5.5.5 avec des bases qu'on notera de la façon suivante : $(s_1, s_2, s_3, s_4, s_5, s_6)$ pour $M_{\rho'_5}$ et $(s_7, s_8, s_9, s_{10}, s_{11}, s_{12})$ pour $M_{\varepsilon\rho'_5}$ où les s_i, $1 \leqq i \leqq 12$, sont homogènes de degrés (5, 9, 13, 17, 21, 25) et (7, 11, 15, 15, 19, 23) respectivement (§ 5.3.5).

Nous allons montrer que le $\mathscr{O}_{\mathscr{U}_k}$-module $\mathscr{F}_{\rho_5} |_{\mathscr{U}_k}$ est isomorphe au module libre de rang 6, L_k, engendré par :

$$\begin{array}{lll} s_1, s_2, s_3, s_7, s_8, s_9 & \text{si} & 0 \leqq k \leqq 4, \\ s_1, s_2, s_3, s_7, s_8, s_{10} & \text{si} & 5 \leqq k \leqq 8. \end{array}$$

Notons c_i la colonne correspondante à $T^-_{\rho_5}(s_i)$, $1 \leqq i \leqq 12$,de la matrice associée à $T^-_{\rho_5}$ avec les bases précédentes. Par abus de notation on écrira, dans ce paragraphe, X, Y, Z, à la place de $q^* X$, $q^* Y$, $q^* Z$, respectivement. En vue de la définition du morphisme de désingularisation q (§ 6.1.5) on a :

Sur $\mathscr{U}_0$, $\mathscr{U}_1$: X divise c_4, c_5, c_6, c_{10}, et Y divise c_2, c_3, d'où on peut éliminer s_4, s_5, s_6, s_{10} et s_{11}, s_{12} respectivement.

Sur $\mathscr{U}_2$, $\mathscr{U}_3$, $\mathscr{U}_4$: X divise c_4, c_5, c_6, et Y divise c_2, c_3; d'où on peut éliminer s_4, s_5, s_6 et s_{11}, s_{12} respectivement. Par ailleurs Z^2 divise c_{10}, donc on peut former une combinaison linéaire de la forme $c_{10}/Z^2 + \star c_5/X$ qui soit divisible par X/Z^2, d'où on peut éliminer s_{10}.

Sur $\mathscr{U}_5$, $\mathscr{U}_6$: par les mêmes propriétés de divisibilité que pour $\mathscr{U}_2$, $\mathscr{U}_3$, $\mathscr{U}_4$, on peut éliminer s_4, s_5, s_6, s_{11}, s_{12}. Or la combinaison linéaire $c_1/Z^2 + \star c_5/X$ précédente est divisible par Z^3/X, d'où on peut éliminer s_9.

Sur $\mathscr{U}_7$: Z divise c_1, c_7 et Z^2 divise c_2, c_3, c_8, c_9; d'où on peut éliminer s_{12}, s_6 et s_9, s_{11}, s_4, s_5 respectivement.

Sur $\mathscr{U}_8$: Y divise c_2, c_3, c_8, c_9 et YZ divise c_4, d'où on peut éliminer s_{11}, s_{12}, s_5, s_6 et s_9 respectivement. Par ailleurs Z^2 divise c_{10}, donc on peut former une combinaison linéaire de la forme $c_{10}/Z^2 + \star c_8/Y$ qui soit divisible par Z^2/Y, d'où on peut éliminer s_9.

La discussion précédente montre que la projection $\varpi_{\rho_5}|_{\mathscr{U}_5}$ induit un homomorphisme surjectif de L_k sur $\mathscr{F}_{\rho_5}|_{\mathscr{U}_k}$; l'injectivité résulte du corollaire 2.10(ii).

Finalement, on a $\deg \mathscr{F}_{\rho_5}|_{d_k}=0$ si $k\neq 5$, $1\leqq k\leqq 8$, car les changements de base sur $\mathscr{U}_{k-1}\cap\mathscr{U}_k$ $(k\neq 5,\ 1\leqq k\leqq 7)$ et sur $\mathscr{U}_8\cap\mathscr{U}_5$ sont triviaux; et $\deg \mathscr{F}_{\rho_5}|_{d_5}=1$, parce que sur $\mathscr{U}_4\cap\mathscr{U}_5$ on a la relation :

$$s_9=\star(v_5+1)\,s_{10}+f(s_1,\ s_2,\ s_3,\ s_7,\ s_8)$$

où f est une combinaison linéaire à coefficients dans $\mathscr{O}_{\mathscr{U}_5}$, et où (v_5+1) provient de la restriction à $\mathscr{U}_5$ du quotient X^2/Z^5.

6.2.4. *Fin de la démonstration.* — La représentation canonique c est (avec les notations de paragraphe 3.7) $\rho_1\oplus\rho_n$ pour le type $\mathbf{A}_n (n\geqq 1)$, ρ_2 pour le type $\mathbf{D}_n (n\geqq 4)$, et ρ_1 pour les types $\mathbf{E}_6$, $\mathbf{E}_7$, $\mathbf{E}_8$. Avec les notations de paragraphe 6.1 pour les composantes irréductibles du diviseur D on trouve dans la démonstration de (i) :

$$\pi(c)=\begin{cases}(2,\ b_{d_1}+b_{d_n}) \text{ pour le type } \mathbf{A}_n & (n\geqq 1),\\ (2,\ b_{d_2}) \text{ pour le type } \mathbf{D}_n & (n\geqq 4),\\ (2,\ b_{d_1}) \text{ pour les types } \mathbf{E}_6,\ \mathbf{E}_7,\ \mathbf{E}_8.\end{cases}$$

Par ailleurs, le cycle fondamental $\mathscr{Z}$ est égal à $\sum_{k=1}^{n} d_k$ pour le type $\mathbf{A}_n$, et il est représenté par les diagrammes suivants pour les autres types :

Type	$\mathscr{Z}$
$\mathbf{D}_n (n\geqq 4)$	$\begin{array}{cccc} & & & 1\\ 1 & 2\ldots & 2 & \\ & & & 1\end{array}$
$\mathbf{E}_6$	$\begin{array}{ccccc}1&2&3&2&1\\ &&2&&\end{array}$
$\mathbf{E}_7$	$\begin{array}{cccccc}2&3&4&3&2&1\\ &&2&&&\end{array}$
$\mathbf{E}_8$	$\begin{array}{ccccccc}2&3&4&5&6&4&2\\ &&&&3&&\end{array}$

On vérifie donc, avec les conventions de paragraphe 1.1, les égalités dans $K(\mathscr{S})$:

$$(\mathscr{Z})=\begin{cases}(0,\ -b_{d_1}-b_{d_n}) \text{ pour le type } \mathbf{A}_n & (n\geqq 1),\\ (0,\ -b_{d_2}) \text{ pour le type } \mathbf{D}_n & (n\geqq 4),\\ (0,\ -b_{d_1}) \text{ pour les types } \mathbf{E}_6,\ \mathbf{E}_7,\ \mathbf{E}_8.\end{cases}$$

Par conséquent on a (dans tous les cas) l'égalité :

$$\pi(2-c)=[\mathscr{Z}].$$

Par ailleurs, on a, pour $i\neq 0$:

$$(c-2)\,\rho_i=\sum_j a_{ij}\,\rho_j-2\,\rho_i$$

d'où :

$$\pi((c-2)\,\rho_i)=(0,\ \sum_{j>0} a_{ij}\,\rho_j-2\,\rho_i).$$

Or l'élément e_{d_i} de $\mathrm{Pic}_{\mathrm{D}}(\tilde{\mathscr{S}})$ s'exprime dans la base $\{b_d\}$ de $\mathrm{Pic}(\tilde{\mathscr{S}})$ de la façon suivante :

$$e_{d_i}=\sum_{j>0} a_{ij}\,b_{d_j}-2\,b_{d_i}$$

où on a d_i à la place de d_{ρ_i}. Par suite de ce qui précède et de la proposition 1.2(iii), on a :

$$\pi((c-2)\,\rho_i)=[d_{\rho_i}],\qquad \forall\, i>0,$$

d'où l'assertion (ii).

L'assertion (i) et le fait que $\pi(\rho_0)=(1,0)$ impliquent (iii). L'assertion (iv) résulte immédiatement de la dernière partie de (ii). L'assertion (v) est conséquence du fait que $\mathrm{R(G)}/(2-c)\,\mathrm{R(G)}$ est isomorphe à $\tilde{\mathrm{K}}^{\mathrm{G}}(\mathscr{V}-\{0\})$ (*voir* § 1.5), et du fait que

$$q^{-1}\circ p|_{\mathscr{V}-\{0\}} : \mathscr{V}-\{0\}\to\tilde{\mathscr{S}}-\mathrm{D}$$

est un morphisme étale, ce qui achève la démonstration du théorème.

7. Corollaire : Caractérisation de la désingularisation minimale

Soit $\mathscr{V}$ *le germe de* $\mathbb{C}^2$ *à l'origine,* G *un sous-groupe fini de* $\mathrm{SL}(2,\mathbb{C})$ *qui agit librement sur* $\mathscr{V}-\{0\}$; $p:\mathscr{V}\to\mathscr{S}=\mathscr{V}/\mathrm{G}$ *le morphisme quotient. Alors la désingularisation minimale* $q:\tilde{\mathscr{S}}\to\mathscr{S}$ *est le platificateur normal minimal de* $\mathscr{V}$ *sur* $\mathscr{S}$.

Démonstration. — Soit $\bar{q}:\mathscr{P}\to\mathscr{S}$ le platificateur normal minimal de $\mathscr{V}$ sur $\mathscr{S}$. En vertu de la proposition 2.4 on sait que $q:\tilde{\mathscr{S}}\to\mathscr{S}$ est un platificateur de $\mathscr{V}$ sur $\mathscr{S}$, qui de plus est normal car $\tilde{\mathscr{S}}$ est lisse. Par suite, le morphisme q se factorise par $\mathscr{P}$, i. e. on a un diagramme commutatif :

$$\begin{tikzcd}
\tilde{\mathscr{S}} \arrow[dr, "\sigma"] \arrow[dd, "q"'] & \\
& \mathscr{P} \arrow[dl, "\bar{q}"] \\
\mathscr{S} &
\end{tikzcd}$$

Par minimalité, le morphisme σ est un isomorphisme en dehors du diviseur exceptionnel $\mathrm{D}\subset\tilde{\mathscr{S}}$. Nous allons montrer que, en fait, σ est un isomorphisme. S'il ne l'est pas, alors, en vertu de la normalité de $\mathscr{P}$ il contracte un sous-diviseur $\mathrm{D}^c\subset\mathrm{D}$, i. e. $\sigma(\mathrm{D}^c)$ est de dimension zéro. Soit ρ une représentation irréductible de G dans un $\mathbb{C}$-espace vectoriel E_ρ. Posons $\overline{\mathscr{F}}_\rho=\bar{q}^*\mathscr{M}_\rho/\mathrm{Tors}_{\mathscr{P}}$ où $\mathscr{M}_\rho=\mathrm{Hom}_{\mathbb{C}[\mathrm{G}]}(\mathrm{E}_\rho,\mathscr{O}_{\mathscr{V}})$; on a :

$$\mathscr{O}_{\mathscr{V}}=\bigoplus_{\rho}(\mathscr{M}_\rho\otimes_{\mathbb{C}}\mathrm{E}_\rho)\qquad\text{d'où}\quad \bar{q}^*\mathscr{O}_{\mathscr{V}}/\mathrm{Tors}_{\mathscr{P}}=\bigoplus_{\rho}(\overline{\mathscr{F}}_\rho\otimes_{\mathbb{C}}\mathrm{E}_\rho).$$

Par conséquent, $\overline{\mathscr{F}}_{\rho}$ est localement libre car $\mathscr{P}$ est un platificateur de $\mathscr{V}$ sur $\mathscr{S}$. De même, le $\mathscr{O}_{\tilde{\mathscr{S}}}$-module $\mathscr{F}_{\rho}=q^{*}\mathscr{M}_{\rho}/\mathrm{Tors}_{\tilde{\mathscr{S}}}$ est localement libre, et par la commutativité du diagramme précédent on a :

$$\sigma^{*}\overline{\mathscr{F}}_{\rho}=\mathscr{F}_{\rho}.$$

Soit ρ telle que la composante irréductible d_{ρ} de D (th. 2.2) soit contractée par σ, i. e. telle que $d_{\rho}\subset D^{c}$. Alors on a une contradiction, car d'une part la restriction de $\mathscr{F}_{\rho}$ à d_{ρ} est non triviale [car $(c_1(\mathscr{F}_{\rho}).d_{\rho})=1$], et d'autre part $\overline{\mathscr{F}}_{\rho}$ est libre au voisinage du point $\sigma(d_{\rho})\in\mathscr{P}$, et par suite $\sigma^{*}\overline{\mathscr{F}}_{\rho}$ est trivial au voisinage de d_{ρ}. Par conséquent D^{c} est vide et σ est un isomorphisme.

RÉFÉRENCES

[A] M. Artin, *On Isolated Rational Singularities of Surfaces* (*Amer. J. Math.*, 88, 1966, p. 129-136).

[Bo] N. Bourbaki, *Groupes et algèbres de Lie*, chap. IV, V, VI, Hermann, Paris, 1968.

[Br] E. Brieskorn, *Rationale Singularitäten Komplexer Flächen* (*Inv. Math.*, 4, 1968, p. 336-358).

[Ch] C. Chevalley, *Invariants of Finite Groups Generated by Reflections* (*Amer. J. Math.*, Vol. 77, 1955, p. 778-782).

[Co] A. M. Cohen, *Finite Complex Reflection Groups* (*Ann. scient. Ec. Norm. Sup.*, 1976, p. 379-436).

[D] P. Duval, *Homographies, Quaternions and Rotations*, Oxford Clarendon Press, 1964.

[E] Euclide, *Éléments*, Livre XIII, § 18.A, Dover, 1956.

[G] G. Gonzalez-Sprinberg, *Résolution de Nash des points doubles rationnels* (*Ann. de l'Inst. Fourier*, vol. 32, n° 2, 1982, p. 111-178).

[G-V] G. Gonzalez-Sprinberg et J. L. Verdier, *Points doubles rationnels et représentations de groupes* (*C.R. Acad. Sc.*, Paris, t. 293, 1981, p. 111-113).

[Gr] A. Grothendieck, *Éléments de géométrie algébrique IV*, Publications Mathématiques de l'I.H.E.S.

[H] F. Hirzebruch, *Uber vierdimensionale Riemannsche Flächen mehrdeutiger analytischer Funktionen von zwei komplexen Veränderlichen* (*Math. Ann.*, vol. 126, 1953, p. 1-22).

[K] F. Klein, *The Icosahedron and the General 5th Degree Equation*, 1884, Dover reprint 1956.

[Kn] H. Knorrer, *Lettre* (décembre 1981).

[L] H. Laufer, *Normal Two-Dimensional Singularities* (*Annals of Math. Studies*, n° 71, Princeton, 1971).

[M] J. McKay, *Graphs, Singularities and Finite Groups* (*Proc. of Symp. in Pure Math.*, vol. 37, 1980, p. 183-186).

[P] Platon, *Timée*.

[Se] J.-P. Serre, *Représentations linéaires des groupes finis*, Hermann, Paris, 1967.

[Sl] P. Slodowy, *Simple Singularities and Simple Algebraic Groups* (Springer *Lecture Notes*, n° 815, 1980).

[St] J. Steinberg, *Subgroups of SU_2 and Dynkin Diagrams*, preprint.

[S-T] G. C. Shephard et J. A. Todd, *Finite Unitary Reflection Groups* (*Canad. J. Math.*, vol. 5, 1953, p. 364-383).

(Manuscrit reçu le 30 avril 1982,
révisé le 25 mars 1983)

G. Gonzalez-Sprinberg
Faculté des Sciences de Tours,
Département de Mathématiques
et
École Normale Supérieure,
Centre de Mathématiques,
E.R.A. 589.
J. L. Verdier
École Normale Supérieure,
Centre de Mathématiques,
E.R.A. 589.

Introduction to the monster Lie algebra.

Groups, Combinatorics, and geometry, p. 99–107, edited by M. W. Liebeck and J. Saxl, L.M.S. lecture note series 165, C.U.P. 1992.

Richard E. Borcherds,

I would like to thank J. M. E. Hyland for suggesting many improvements to this paper.

The monster sporadic simple group, of order

$$2^{46}3^{20}5^{9}7^{6}11^{2}13^{3}17.19.23.29.31.41.47.59.71,$$

acts on a graded vector space $V=\oplus_{n\in\mathbf{Z}}V_n$ constructed by Frenkel, Lepowsky and Meurman [Fre] The dimension of V_n is equal to the coefficient $c(n)$ of the elliptic modular function

$$j(\tau)-744=\sum_n c(n)q^n=q^{-1}+196884q+21493760q^2+\cdots$$

(where we write q for $e^{2\pi i\tau}$, and $\mathrm{Im}(\tau)>0$). The main problem is to describe V as a graded representation of the monster, or in other words to calculate the trace $\mathrm{Tr}(g|V_n)$ of each element g of the monster on each space V_n. The best way to describe this information is to define the Thompson series

$$T_g(q)=\sum_{n\in\mathbf{Z}}\mathrm{Tr}(g|V_n)q^n$$

for each element g of the monster, so that our problem is to work out what these Thompson series are. For example, if 1 is the identity element of the monster then $\mathrm{Tr}(1|V_n)=\dim(V_n)=c(n)$, so that the Thompson series $T_1(q)=j(\tau)-744$ is the elliptic modular function. McKay, Thompson, Conway and Norton conjectured [Con] that the Thompson series $T_g(q)$ are always Hauptmoduls for certain modular groups of genus 0 (I will explain what this means in a moment.) In this paper we describe the proof of this in [Bor] using an infinite dimensional Lie algebra acted on by the monster called the monster Lie algebra. These Hauptmoduls are known explicitly, so this gives a complete description of V as a representation of the monster.

We now recall the definition of a Hauptmodul. The group $SL_2(\mathbf{Z})$ acts on the upper half plane $H=\{\tau\in\mathbf{C}|\mathrm{Im}(\tau)>0\}$ by $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)(\tau)=\frac{a\tau+b}{c\tau+d}$. The elliptic modular function $j(\tau)$ is more or less the simplest function defined on H that is invariant under $SL_2(\mathbf{Z})$ in much the same way that the function $e^{2\pi i\tau}$ is the simplest function invariant under $\tau\mapsto\tau+1$. The element $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ of $SL_2(\mathbf{Z})$ takes τ to $\tau+1$, so in particular $j(\tau)$ is periodic and can be written as a Laurent series in $q=e^{2\pi i\tau}$. The exact expression for j is

$$j(\tau)=\frac{(1+240\sum_{n>0}\sigma_3(n)q^n)^3}{q\prod_{n>0}(1-q^n)^{24}}$$

where $\sigma_3(n)=\sum_{d|n}d^3$ is the sum of the cubes of the divisors of n; see any book on modular forms or elliptic functions, for example [Ser]. Another way of thinking about j is that it is an isomorphism from the quotient space $H/SL_2(\mathbf{Z})$ to the complex plane, which can be thought of as the Riemann sphere minus the point at infinity.

There is nothing particularly special about the group $SL_2(\mathbf{Z})$ acting on H. We can consider functions invariant under any group G acting on H, for example some group G of finite index in $SL_2(\mathbf{Z})$. If G satisfies some mild conditions then the quotient H/G will again be a compact Riemann surface with a finite number of points removed. If this compact Riemann surface is a sphere, rather than something of higher genus, then we say that G is a genus 0 group. When this happens there is a more or less unique map from H/G to the complex plane $\mathbf{C}$, which induces a function from H to $\mathbf{C}$ which is invariant under the group G acting on H. When this function is correctly normalized it is called a Hauptmodul for the genus 0 group G. For example, $j(\tau) - 744$ is the Hauptmodul for the genus 0 group $SL_2(\mathbf{Z})$.

For example, we could take G to be the group $\Gamma_0(2)$, where $\Gamma_0(N) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbf{Z}) | c \equiv 0 \bmod N\}$. The quotient H/G is then a sphere with 2 points removed, so that G is a genus 0 group. Its Hauptmodul starts off $T_{2-}(q) = q^{-1} + 276q - 2048q^2 + \cdots$, and is equal to the Thompson series for a certain element of the monster of order 2 (of type 2B in atlas notation). Similarly $\Gamma_0(N)$ is a genus 0 subgroup for several other values of N which correspond to elements of the monster. (However the genus of $\Gamma_0(N)$ tends to infinity as N increases, so there are only a finite number of integers N for which it has genus 0; more generally Thompson has shown that there are only a finite number of conjugacy classes of genus 0 subgroups of $SL_2(\mathbf{R})$ which are commensurable with $SL_2(\mathbf{Z})$.) The first person to notice any connection between the monster and genus 0 subgroups of $SL_2(\mathbf{R})$ was probably Ogg, who observed that the primes p for which the normalizer of $\Gamma_0(p)$ in $SL_2(\mathbf{R})$ has genus 0 are exactly those which divide the order of the monster.

So the problem we want to solve is to calculate the Thompson series $T_g(\tau)$ and show that they are Hauptmoduls of genus 0 subgroups of $SL_2(\mathbf{R})$. The difficulty with doing this is as follows. Frenkel, Lepowsky, and Meurman constructed V as the sum of two subspaces V^+ and V^-, which are the $+1$ and -1 eigenspaces of an element of order 2 in the monster. If an element g of the monster commutes with this element of order 2, then it is not difficult to to work out its Thompson series $T_g(q) = \sum_n \mathrm{Tr}(g|V_n)q^n$ as the sum of two series given by its traces on V^+ and V^- and it would be tedious but straightforward to check that these all gave Hauptmoduls. However if an element of the monster is not conjugate to something that commutes with this element of order 2 then there is no obvious direct way of working out its Thompson series, because it muddles up V^+ and V^- in a very complicated way.

We now give a bird's eye view of the rest of this paper. We first construct a Lie algebra from V, called the monster Lie algebra. This is a generalized Kac-Moody algebra, so we recall some facts about such algebras, and in particular show how each such algebra gives an identity called its denominator formula. A well known example of this is that the denominator formulas of affine Kac-Moody algebras are the Macdonald identities. The denominator formula for the monster Lie algebra turns out to be the product formula for the elliptic modular function j, and we use this to work out the structure of the monster Lie algebra, and in particular to find its "simple roots". Once we know the simple roots, we can write down a sort of twisted denominator formula for each element of the monster group. These twisted denominator formulas imply some relations between the coefficients of the Thompson series of the monster, which are strong enough to characterize them and verify that they are indeed Hauptmoduls.

We now introduce the the monster Lie algebra. This is a $\mathbf{Z}^2$-graded Lie algebra, whose piece of degree $(m,n) \in \mathbf{Z}^2$ is isomorphic as a module over the monster to V_{mn} if $(m,n) \neq (0,0)$ and to $\mathbf{R}^2$ if $(m,n) = (0,0)$, so for small degrees it looks something like

$$\begin{array}{ccccccccc}
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & 0 & 0 & 0 & 0 & V_3 & V_6 & V_9 & \cdots \\
\cdots & 0 & 0 & 0 & 0 & V_2 & V_4 & V_6 & \cdots \\
\cdots & 0 & 0 & V_{-1} & 0 & V_1 & V_2 & V_3 & \cdots \\
\cdots & 0 & 0 & 0 & \mathbf{R}^2 & 0 & 0 & 0 & \cdots. \\
\cdots & V_3 & V_2 & V_1 & 0 & V_{-1} & 0 & 0 & \cdots \\
\cdots & V_6 & V_4 & V_2 & 0 & 0 & 0 & 0 & \cdots \\
\cdots & V_9 & V_6 & V_3 & 0 & 0 & 0 & 0 & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

This is a very big Lie algebra because the dimension of V_n increases exponentially fast; most infinite dimensional Lie algebras that occur in mathematics can be graded so that the dimensions of the graded pieces have only polynomial growth.

Very briefly, the construction of this Lie algebra goes as follows (this construction is not used later, so it does not matter if the reader finds this meaningless). The graded vector space V is a vertex algebra; roughly speaking this means that V has an infinite number of products on it which satisfy some rather complicated identities. (The Griess product on the 196884 dimensional space V_1 is a special case of one of the vertex algebra products.) We can also construct a vertex algebra from any even lattice, and the tensor product of two vertex algebras is a vertex algebra. We take the tensor product of the monster vertex algebra V with that of the two dimensional even Lorentzian lattice (defined by the matrix $\left(\begin{smallmatrix}0,1\\1,0\end{smallmatrix}\right)$) to get a new vertex algebra W. The monster Lie algebra is then the space of physical states of the vertex algebra W; this is a subquotient of W, which is roughly a space of highest weight vectors for an action of the Virasoro algebra on W. The subquotient can be identified exactly using the no-ghost theorem from string theory, and turns out to be as described above.

We need to know what the structure of the monster Lie algebra is. It turns out to be something called a generalized Kac-Moody algebra, so we will now have a digression to explain what these are.

To motivate Kac-Moody algebras we first look at the structure of finite dimensional simple Lie algebras. A typical finite dimensional Lie algebra is $sl_4(\mathbf{R})$, the algebra of 4×4 matrices of trace 0 with bracket given by $[a,b] = ab - ba$. If we look at this Lie algebra we see that it has three subalgebras $sl_2(\mathbf{R})$ as indicated below (as well as many others):

$$\begin{pmatrix} * & * & & \\ * & * & & \\ & & & \\ & & & \end{pmatrix}, \begin{pmatrix} & & & \\ & * & * & \\ & * & * & \\ & & & \end{pmatrix}, \begin{pmatrix} & & & \\ & & & \\ & & * & * \\ & & * & * \end{pmatrix}.$$

The Dynkin diagram of $sl_4(\mathbf{R})$ is constructed by drawing a point for each of these sl_2's, and connecting them by various numbers of lines depending on how the sl_2's are arranged; for example, if the two sl_2's commute then we draw no lines between the corresponding points, and if the sl_2's overlap in one corner we draw a single line between the points. The Dynkin diagram of sl_4 is therefore $\bullet - \bullet - \bullet$, where the two outer dots are not joined because the corresponding sl_2's of sl_4 commute, and so on. Conversely we can reconstruct sl_4 from its Dynkin diagram by writing down some generators and relations from the Dynkin diagram. Roughly speaking, sl_4 is generated by an sl_2 for each point of the Dynkin diagram together with some relations depending on the lines between points; for example, if two points are not joined we add relations saying that the corresponding sl_2's commute. We can construct all other finite dimensional split semisimple Lie algebras from their Dynkin diagrams (usually denoted by A_n, B_n, C_n, D_n, E_6, E_7, E_8, F_4, and G_2) in a similar way. If we are given some graph which is not one of these Dynkin diagrams of finite dimensional Lie algebras we can still write down generators and relations; the main difference is that these Lie algebras will be infinite dimensional rather than finite dimensional. We can think of a Kac-Moody algebra as a Lie algebra generated by a copy of sl_2 for each point of its Dynkin diagram.

The simplest example of an infinite dimensional Kac-Moody algebra is the Lie algebra $sl_2(\mathbf{R}[z,z^{-1}])$ of 2×2 matrices whose entries are Laurent polynomials. (This is a slight simplification; the Kac-Moody algebra is really a central extension of this Lie algebra, but we will ignore this.) The Dynkin diagram is $\bullet = \bullet$, so this algebra is generated by 2 copies of $sl_2(\mathbf{R})$, which are the subalgebras of elements of the form $\left(\begin{smallmatrix} a,b \\ c,-a \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} a,bz \\ cz^{-1},-a \end{smallmatrix}\right)$. The algebra $sl_2(\mathbf{R}[z,z^{-1}])$ can be $\mathbf{Z}^2$-graded by letting $\left(\begin{smallmatrix} 0,z^n \\ 0,0 \end{smallmatrix}\right)$ have degree $(2n+1,1)$, letting $\left(\begin{smallmatrix} z^n,0 \\ 0,-z^n \end{smallmatrix}\right)$ have degree (2n,0), and letting $\left(\begin{smallmatrix} 0,0 \\ z^n,0 \end{smallmatrix}\right)$ have degree $(2n-1,-1)$. The picture we get looks like

$$\begin{array}{ccccccccc}
\cdots\begin{pmatrix} 0 & 0 \\ z^{-2} & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 0 \\ z^{-1} & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 0 \\ z & 0 \end{pmatrix}\cdots \\
\cdots & \begin{pmatrix} z^{-1} & 0 \\ 0 & -z^{-1} \end{pmatrix} & & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & & \begin{pmatrix} z & 0 \\ 0 & -z \end{pmatrix} & \cdots \\
\cdots\begin{pmatrix} 0 & z^{-1} \\ 0 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & z^2 \\ 0 & 0 \end{pmatrix}\cdots
\end{array}$$

where on each point of $\mathbf{Z}^2$ we have written something that spans the corresponding subspace of the Lie algebra.

We now describe the denominator formula of Kac-Moody algebras, and as an example show that for the Lie algebra $sl_2(\mathbf{R}[z,z^{-1}])$ it becomes the Jacobi triple product identity. The characters $Ch(V)$ of finite dimensional irreducible representations V of finite dimensional simple Lie algebras are described by the Weyl character formula

$$Ch(V) = \frac{\sum_{w\in W}\det(w)e^{w(\lambda-\rho)}}{e^{-\rho}\prod_{\alpha>0}(1-e^{\alpha})^{\mathrm{mult}(\alpha)}}$$

and the same formula turns out to be true for the characters of lowest weight representations of Kac-Moody algebras. The only case we will use this is when V is the trivial one dimensional representation with character 1, when the character formula becomes the denominator formula

$$e^{-\rho}\prod_{\alpha>0}(1-e^{\alpha})^{\mathrm{mult}(\alpha)}=\sum_{w\in W}\det(w)e^{-w(\rho)}.$$

Rather than explain precisely what all the terms in this formula mean, we will just describe them in the case of $sl_2(\mathbf{R}[z,z^{-1}])$. The product is over the "positive roots" α; for $sl_2(\mathbf{R}[z,z^{-1}])$ these are the vectors $\alpha=(m,n)$ with $m>0$ for which there is something in $sl_2(\mathbf{R}[z,z^{-1}])$ of that degree, in other words the vectors $(2m,0)$, $(2m-1,1)$, and $(2m-1,-1)$ for $m>0$. The multiplicity $\mathrm{mult}(\alpha)$ is the dimension of the subspace of that degree, which for $sl_2(\mathbf{R}[z,z^{-1}])$ is always 1. If we put $e^{(m,n)}=x^m y^n$, then the product becomes

$$\prod_{m>0}(1-x^{2m})(1-x^{2m-1}y)(1-x^{2m-1}y^{-1}).$$

The sum is a sum over all elements of the "Weyl group" W, which is a reflection group generated by one reflection for each point of the Dynkin diagram. In this case the Weyl group is the infinite dihedral group, which has an infinite cyclic subgroup of index 2, so the sum over the Weyl group is essentially a sum over the integers $\mathbf{Z}$. If we work out explicitly what it is, the Weyl denominator formula becomes

$$\prod_{m>0}(1-x^{2m})(1-x^{2m-1}y)(1-x^{2m-1}y^{-1})=\sum_{n\in\mathbf{Z}}(-1)^n x^{n^2}y^n.$$

which is the Jacobi triple product identity.

Not surprisingly, we can do the same with sl_2 replaced by any other simple finite dimensional Lie algebra, and we then obtain some identities called the Macdonald identities. (In fact there is more than one Macdonald identity for some finite dimensional Lie algebras, because there are sometimes ways of "twisting" this construction.)

Generalized Kac-Moody algebras are rather like Kac-Moody algebras except that we are allowed to glue together the sl_2's in more complicated ways, and are also allowed to use Heisenberg Lie algebras as well as sl_2's to generate the algebra. In technical terms, the roots of a Kac-Moody algebra may be either real (norm >0) or imaginary (norm ≤ 0) but all the simple roots must be real. Generalized Kac-Moody algebras may also have imaginary simple roots. They have a denominator formula which is similar to the one above, except that it has some extra correction terms corresponding to the imaginary simple roots. (The simple roots of a generalized Kac-Moody algebra correspond to a minimal set of generators for the subalgebra corresponding to the positive roots; for example, the simple roots of $sl_2(\mathbf{R}[z,z^{-1}])$ are $(1,1)$ and $(1,-1)$. For Kac-Moody algebras the simple roots also correspond to the points of the Dynkin diagram and to the generators of the Weyl group.)

We now return to the monster Lie algebra. This is a generalized Kac-Moody algebra, and we are going to write down its denominator formula, which says that a product over the positive roots is a sum over the Weyl group. The positive roots are the vectors (m,n) with $m>0$, $n>0$, and the vector $(1,-1)$, and the root (m,n) has multiplicity $c(mn)$. The Weyl group has order 2 and its nontrivial element maps (m,n) to (n,m), so it exchanges

$p = e^{(1,0)}$ and $q = e^{(0,1)}$. The denominator formula is the product formula for the j function

$$p^{-1}\prod_{m>0,n\in\mathbf{Z}}(1-p^mq^n)^{c(mn)} = j(p)-j(q).$$

(The left side is apparently not antisymmetric in p and q but in fact it is because of the factor of $p^{-1}(1-p^1q^{-1})$ in the product.) The reason why we get $j(p)$ and $j(q)$ rather than just a monomial in p and q on the right hand side (as we would for ordinary Kac-Moody algebras) is because of the correction due to the imaginary simple roots of M. The simple roots of M correspond to a set of generators of the subalgebra E of the elements of M whose degree is to the right of the y axis (so the roots of E are the positive roots of M), and turn out to be the vectors $(1,n)$ each with multiplicity $c(n)$. In the picture of the monster Lie algebra given earlier, the simple roots are given by the column just to the right of the one containing $\mathbf{R}^2$. The sum of the simple root spaces is isomorphic to the space V. The simple root $(1,-1)$ is real of norm 2, and the simple roots $(1,n)$ for $n>0$ are imaginary of norm $-2n$ and have multiplicity $c(n)=\dim(V_n)$. As these multiplicities are exactly the coefficients of the j function, it is not surprising that j appears in the correction caused by the imaginary simple roots. This discussion is slightly misleading because we have implied that we obtain the product formula of the j function as the denominator formula of the monster Lie algebra by using our knowledge of the simple roots; in fact we really have to use this argument in reverse, using the product formula for the j function in order to work out what the simple roots of the monster Lie algebra are.

We will now extract information about the coefficients of the Thompson series $T_g(\tau)$ from a twisted denominator formula for the monster Lie algebra. For an arbitrary generalized Kac-Moody algebra there is a more high powered version of the Weyl denominator formula which states that

$$\Lambda(E) = H_*(E),$$

where E is the subalgebra corresponding to the positive roots. Here $\Lambda(E) = \Lambda^0(E)\ominus\Lambda^1(E)\oplus\Lambda^2(E)\ldots$ is the alternating sum of the exterior powers of E, and similarly $H_*(E)$ is the alternating sum of the homology groups $H_i(E)$ of the Lie algebra E (see [Car]). This identity is true for any Lie algebra because the H_i's are the homology groups of a complex whose terms are the Λ^i's. The left hand side corresponds to a product over the positive roots because $\Lambda(A\oplus B)=\Lambda(A)\otimes\Lambda(B)$, $\Lambda(A)=1-A$ if A is one dimensional, and E is the sum of $\mathrm{mult}(\alpha)$ one dimensional spaces for each positive root α. It is more difficult to identify H_* with a sum over the Weyl group, and we do this roughly as follows. For Kac-Moody algebras H_i turns out to have dimension equal to the number of elements in the Weyl group of length i; for finite dimensional Lie algebras this was first observed by Bott, and was used by Kostant to give a homological proof of the Weyl character formula. The sum over the homology groups can therefore be identified with a sum over the Weyl group. For generalized Kac-Moody algebras things are a bit more complicated. The sum over the homology groups becomes a sum over the Weyl group, which is generated by an involution for each simple root, and the things we sum contain terms corresponding to the imaginary simple roots.

Anyway, we can work out the homology groups of E explicitly provided we know the simple roots of our Lie algebra; for example, the first homology group H_1 is the sum of the simple root spaces. For the monster Lie algebra we have worked out the simple roots using its denominator formula, which is the product formula for the j function, and the homology groups of E turn out to be $H_0 = \mathbf{R}$, $H_1 = \sum_{n\in\mathbf{Z}} V_n p q^n$, $H_2 = \sum_{m>0} V_m p^{m+1}$, and all the higher homology groups are 0. Each homology group is a $\mathbf{Z}^2$-graded representation of the monster, and we use the p's and q's to keep track of the grading. If we substitute these values into the formula $\Lambda(E) = H_*$ we find that

$$\Lambda(\sum_{n\in\mathbf{Z},m>0} V_{mn} p^m q^n) = \sum_m V_m p^{m+1} - \sum_n V_n p q^n.$$

Both sides of this are virtual graded representations of the monster. If we replace everything by its dimension we recover the product formula for the j function. More generally, we can take the trace of some element of the monster on both sides, which after some calculation gives the identity

$$p^{-1}\exp\Big(-\sum_{i>0}\sum_{m>0,n\in\mathbf{Z}} \mathrm{Tr}(g^i|V_{mn}) p^{mi} q^{ni}/i\Big) = \sum_{m\in\mathbf{Z}} \mathrm{Tr}(g|V_m) p^m - \sum_{n\in\mathbf{Z}} \mathrm{Tr}(g|V_n) q^n$$

where $\mathrm{Tr}(g|V_n)$ is the trace of g on the vector space V_n.

These relations between the coefficients of the Thompson series turn out to be strong enough to characterize them and check that they are Hauptmoduls for genus 0 subgroups. Unfortunately this final step of the proof is rather messy. The relations above determine the Thompson series from their first few coefficients. Norton and Koike checked that certain modular functions of genus 0 also satisfy the same recursion relations. We can therefore prove that the Thompson series $T_g(q)$ are modular functions of genus 0 by checking that the first few coefficients of both functions are the same. Norton has conjectured that Hauptmoduls with integer coefficients are essentially the same as functions satisfying relations similar to the ones above, and a conceptual proof or explanation of this would be a big improvement to the final step of the proof. (It should be possible to prove this conjecture by a very long and tedious case by case check, because all functions which are either Hauptmoduls or which satisfy the relations above can be listed explicitly.)

Bibliography.

[Bor] R. E. Borcherds, Monstrous moonshine and monstrous Lie superalgebras, Invent. Math. 109, 405-444 (1992).

[Car] H. Cartan, S. Eilenberg, Homological algebra, Princeton University Press 1956.

[Con] J. H. Conway, S. Norton, Monstrous moonshine, Bull. London. Math. Soc. 11 (1979) 308-339.

[Fre] I. B. Frenkel, J. Lepowsky, A. Meurman, Vertex operator algebras and the monster, Academic press 1988.

[Ser] J. P. Serre, A course in arithmetic. Graduate texts in mathematics 7, Springer-Verlag, 1973.